

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 08:57:07 ; Search time 24.95 Seconds

(without alignments)  
1122.233 Million cell updates/sec

Title: US-09-579-383-3

Perfect score: 1982

Sequence: 1 NMFTVKYSFLVTCILLCILLS.....SDNAAHNEQLAIEYEVESLH 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDSR/gcgdata/geneseq/AA1980.DAT:\*
- 2: /SIDSR/gcgdata/geneseq/AA1981.DAT:\*
- 3: /SIDSR/gcgdata/geneseq/AA1982.DAT:\*
- 4: /SIDSR/gcgdata/geneseq/AA1983.DAT:\*
- 5: /SIDSR/gcgdata/geneseq/AA1984.DAT:\*
- 6: /SIDSR/gcgdata/geneseq/AA1985.DAT:\*
- 7: /SIDSR/gcgdata/geneseq/AA1986.DAT:\*
- 8: /SIDSR/gcgdata/geneseq/AA1987.DAT:\*
- 9: /SIDSR/gcgdata/geneseq/AA1988.DAT:\*
- 10: /SIDSR/gcgdata/geneseq/AA1989.DAT:\*
- 11: /SIDSR/gcgdata/geneseq/AA1990.DAT:\*
- 12: /SIDSR/gcgdata/geneseq/AA1991.DAT:\*
- 13: /SIDSR/gcgdata/geneseq/AA1992.DAT:\*
- 14: /SIDSR/gcgdata/geneseq/AA1993.DAT:\*
- 15: /SIDSR/gcgdata/geneseq/AA1994.DAT:\*
- 16: /SIDSR/gcgdata/geneseq/AA1995.DAT:\*
- 17: /SIDSR/gcgdata/geneseq/AA1996.DAT:\*
- 18: /SIDSR/gcgdata/geneseq/AA1997.DAT:\*
- 19: /SIDSR/gcgdata/geneseq/AA1998.DAT:\*
- 20: /SIDSR/gcgdata/geneseq/AA1999.DAT:\*
- 21: /SIDSR/gcgdata/geneseq/AA2000.DAT:\*
- 22: /SIDSR/gcgdata/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1982	100.0	378	22	AA1980.DAT:*
2	635	32.0	587	22	AA1981.DAT:*
3	135.5	6.8	298	22	AA1982.DAT:*
4	135.5	6.8	298	22	AA1983.DAT:*
5	134.5	6.8	509	22	AA1984.DAT:*
6	130	6.6	399	22	AA1985.DAT:*
7	128.5	6.5	361	18	AA1986.DAT:*
8	127.5	6.4	561	16	AA1987.DAT:*
9	127.5	6.4	561	17	AA1988.DAT:*
10	127.5	6.4	561	19	AA1989.DAT:*
11	126	6.4	555	21	AA1990.DAT:*

12	124.5	6.3	424	15	AA1991.DAT:*
13	122.5	6.2	490	21	AA1992.DAT:*
14	120.5	6.1	424	21	AA1993.DAT:*
15	120.5	6.1	424	21	AA1994.DAT:*
16	117.5	5.9	866	17	AA1995.DAT:*
17	117.5	5.9	866	21	AA1996.DAT:*
18	114	5.8	536	21	AA1997.DAT:*
19	113.5	5.7	561	15	AA1998.DAT:*
20	113	5.7	1215	21	AA1999.DAT:*
21	113	5.7	1215	22	AA2000.DAT:*
22	112.5	5.7	649	20	AA2001.DAT:*
23	112.5	5.7	668	20	AA2002.DAT:*
24	111	5.6	461	14	AA2003.DAT:*
25	111	5.6	325	20	AA2004.DAT:*
26	110.5	5.6	366	20	AA2005.DAT:*
27	110.5	5.6	376	20	AA2006.DAT:*
28	110.5	5.6	376	20	AA2007.DAT:*
29	110.5	5.6	376	20	AA2008.DAT:*
30	110.5	5.6	504	19	AA2009.DAT:*
31	108.5	5.5	1453	13	AA2010.DAT:*
32	107.5	5.4	541	19	AA2011.DAT:*
33	107.5	5.4	554	18	AA2012.DAT:*
34	107	5.4	554	21	AA2013.DAT:*
35	105	5.3	684	21	AA2014.DAT:*
36	104.5	5.3	1453	14	AA2015.DAT:*
37	104	5.2	851	21	AA2016.DAT:*
38	104	5.2	946	21	AA2017.DAT:*
39	103	5.2	452	22	AA2018.DAT:*
40	103	5.2	507	18	AA2019.DAT:*
41	103	5.2	1185	19	AA2020.DAT:*
42	102.5	5.2	377	22	AA2021.DAT:*
43	102.5	5.2	398	22	AA2022.DAT:*
44	102.5	5.2	787	22	AA2023.DAT:*
45	102	5.1	227	20	AA2024.DAT:*

## ALIGNMENTS

RESULT 1

ID AAB49985 standard; Protein: 378 AA.

AC AAB49985;

DT 13-MAR-2001 (first entry)

XX

DE P. falciparum chitinase SEQ ID NO: 3.

XX

KW Malaria; mosquito; chitinase; fungal disease; parasitic disease;

XX

OS Plasmodium falciparum.

XX

PN W0200073488-A1.

XX

PD 07-DEC-2000.

XX

PE 26-MAY-2000; 2000WO-US14536.

XX

PR 28-MAY-1999; 99US-0136508.

XX

PR 03-FEB-2000; 2000US-0180051.

XX

(TEXTA ) UNIV TEXAS SYSTEM.

PI Vinetiz JM.

XX

XX WPI: 2001-061553/07.

XX

XX N-PSDB: AAC89668.

XX

XX New nucleic acid encoding a Plasmodium species chitinase is useful for

XX

XX preventing transmission of malaria by mosquito feeding on subject that

XX

XX may harbor Plasmodium species organisms

```
XX Claim 29; Page 124-125; 137pp; English.
PS
XX
CC The present invention provides the protein and coding sequences of the
CC Plasmodium falciparum and P. gallinaceum chitinase enzymes. These
CC organisms are the cause of malaria in humans. The sequences are useful
CC in the prevention and treatment of malaria, fungal diseases, parasitic
CC diseases and veterinary diseases, in preventing the transmission of
CC malaria and in the control of arthropod pests in agriculture.
XX
SQ Sequence 378 AA;

Query Match 100.0%; Score 1982; DB 22; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.1e-177;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNFTVYKSFVLYTCLLCCLSTVSYVTEGHARARQESRKNRREITKTESGKIIGQYYP 60
DB 1 mntfvysfvlvcllccllstlysvvteghararqesrknprelltktesgkllggyyp 60
QY 61 SWVSYHNKLDLPNINVMSFAKMDLSYDIESIVSGPLTFKSLIGLEYIGLNEYFND 120
DB 61 swvsyhnkldlpnlnvmsfakmdlsydiesivsgpltfksllgleyiglneyfnd 120
QY 121 AMNLRRARPDIIMLSLGGETYHPSSFDSALNVEKIANLVDELGFDDIDVDYEPNCSFD 180
DB 121 amnlrrarpdiimlslggetyhpssfdsalnvekianlvdelgfdgidvdyepngsfd 180
QY 181 GUNDRKKAFFQVQYVTKRYMCDKILISQSSNGALSCIGFNDPKKICDDDEAPYNSK 240
DB 181 gundrkkaaffvqvvtkrwymcdkilsqssngalscigfndppkkicdddeapynsk 240
QY 241 YNNKPPVKKELRAAOMASAGAIYLMNNLKMDIMVFOVFNTNSTDSTVMKELDSY 300
DB 241 ynnkppvkkelraaomasagaiylmnnlkdmdimvfovfntnstdstvmkeldsy 300
QY 301 AYGKKYDYVILNGFTLMPESTPFNPNDKMLVNSIGDFVTENKLNKRADGFGMLSSD 360
DB 301 aygkkydyvilngftlmpestpfpnpndkmlvnsigdfvtenklnkradgfgmlssd 360
QY 361 NAAHNEQALIEYFVESLH 378
DB 361 naahneqaleiyfveslh 378

RESULT 2
AAB49986
ID AAB49986 standard; Protein; 587 AA.
XX
AC AAB49986;
XX
DT 13-MAR-2001 (first entry)
XX
DE P. gallinaceum chitinase SEQ ID NO: 4.
XX
KW Malaria; mosquito; chitinase; fungal disease; parasitic disease;
KW veterinary disease; arthropod pest.
XX
OS Plasmodium gallinaceum.
XX
PN WO200073488-A1.
XX
PD 07-DEC-2000.
XX
PE 26-MAY-2000; 2000MO-US14536.
XX
PR 28-MAY-1999; 99US-0136508.
XX
PR 03-FEB-2000; 2000US-0180051.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Vinetz JM;
```

```
XX WPI: 2001-061553/07..
DR N-PSDB; AAC89669.
XX
PT New nucleic acid encoding a Plasmodium species chitinase is useful for
PT preventing transmission of malaria by mosquito feeding on subject that
PT may harbor Plasmodium species organisms
XX
XX Claim 29; Page 125-127; 137pp; English.
PS
XX
CC The present invention provides the protein and coding sequences of the
CC Plasmodium falciparum and P. gallinaceum chitinase enzymes. These
CC organisms are the cause of malaria in humans. The sequences are useful
CC in the prevention and treatment of malaria, fungal diseases, parasitic
CC diseases and veterinary diseases, in preventing the transmission of
CC malaria and in the control of arthropod pests in agriculture.
XX
SQ Sequence 587 AA;

Query Match 32.0%; Score 635; DB 22; Length 587;
Best Local Similarity 37.4%; Pred. No. 7.9e-51;
Matches 129; Conservative 82; Mismatches 126; Indels 8; Gaps 5;

QY 37 RKNPREIITKFKESGKIIGQYPSWVSYNHNK---DLNPNLNVMSFAKMDLSTDSI 93
DB 89 rksprqllieykkrkqgliaqyyswsgdrakhmidsnpwvslylafatimlydvs 148
QY 94 ESTVSGPLPFLSLGLGLEYIGLNEYFNDAMNLRARPDIIMLSLGGETYHPSSFDSALN 153
DB 149 rpfngqrflrlrkghleyetygmnlrlfrtkrvpavillslgety-mildekyd 207
QY 154 VEKIANLVDELGFDDIDVDYEPNCSFDGLNDRKADFEVQYVTKRYMCDKILISQSS 213
DB 208 vdkllkdvndldvgldvephgkfylnelnstnyyiklhlrltklpeklisgsg 267
QY 214 SNGALSCIGFNDPKKICDDDEAPYNSKYFNKP-DYKELLRAAOMASAGAIYLMNNLK 272
DB 268 sngalscv--sgvasfckdeespynkflseqltnkelhrraamlsagflnflntake 325
QY 273 MIDWFOVFNTNSTDSTVMKELDSYAYYGGKYDVIIIMGFTLMPESTPFNPNDKMLV 332
DB 326 kidlvifqtn-lettpdlmvdmylshlyfglkylnlilgfslenrgyfspenkell 384
QY 333 KSIDGFVTENKLNKRADGFGMLSSDNAHNEQALIEYFVESL 377
DB 385 elvgktlhdkngnnrardgigivhlfmkeqjprgsfdvdlfnl 429

RESULT 3
AAY33721
ID AAY33721 standard; Protein; 298 AA.
XX
AC AAY33721;
XX
DT 09-NOV-1999 (first entry)
XX
DE Streptomyces albidoflavus mature chitobiosidase.
XX
KW Chitin; chitinolytic enzyme; insect; fungus; resistance;
KW transgenic plant; insecticide; fungicide.
XX
OS Streptomyces albidoflavus.
XX
PN WO9942594-A1.
XX
PD 26-AUG-1999.
XX
PR 12-FEB-1999; 99MO-US03176.
XX
PR 18-FEB-1998; 98US-0025691.
XX
PA (CORR ) CORNELL RES FOUND INC.
```

XX Broadway RM, Harman GE;  
PI  
XX WPI: 1999-527480/44.  
DR N-PSDB; AA206824.  
XX  
PT New nucleic acid encoding chitinolytic enzyme active under alkaline  
PT conditions, used to generate plants resistant to insects and fungi  
PT and to produce recombinant enzyme  
XX  
XX Claim 6; Page 10-11; 90pp; English.  
XX  
XX This sequence represents Streptomyces albidoflavus chitinobiosidase.  
CC Chitinobiosidase hydrolyses chitin, an insoluble linear  
CC beta-1,4-linked polymer of N-acetyl-beta-D-glucosamine found in all  
CC arthropods, yeast, most fungi, and some stages of nematodes. There are  
CC three types of chitinolytic enzyme activity: glucosaminidase, which  
CC cleaves monomeric units from the terminal end of chitin; chitinobiosidase,  
CC which cleaves dimeric units from the terminal ends; and  
CC endochitinase, which randomly cleaves the chitin molecule internally.  
CC Chitinobiosidase and endochitinase (AA33722) from *S. albidoflavus* are  
CC active under alkaline conditions. This makes them particularly effective  
CC in controlling insects, as these enzymes can be ingested by insects and  
CC can then attack them by degrading their chitin-containing, alkaline  
CC digestive tracts. These chitinolytic enzymes can be applied directly to a  
CC plant to act as an insecticide or fungicide, or non-pathogenic bacteria  
CC transformed with the DNA encoding either or both enzymes can be applied  
CC to a plant. Alternatively, transgenic plants can be produced which  
CC express either or both enzymes in some or all of their tissues. A wide  
CC variety of commercially important crops could benefit from use of these  
CC enzymes e.g, sugar cane, potato, cotton and coffee.  
XX  
XX Sequence 298 AA;  
SQ  
Query Match 6.8%; Score 135.5; DB 20; Length 298;  
Best Local Similarity 21.7%; Pred. No. 0.0002;  
Matches 55; Conservative 50; Mismatches 76; Indels 73; Gaps 12;  
QY 55 IQGYPSWVSYNH-----NLKDLNPNLVNVMSPFAKMDLSYDSISIVGSPLLFKSLGL 109  
DB 10 vlgv---wqnfngatvgtladvdpaydlavsfadatanagel-----tftl 54  
QY 110 EYIGLNEFYNDAMNLRKRPDI-----MLSLGGETYHPSPSDSLANVEKIAN-- 159  
DB 55 dsvglggytde-----qfradlaakqadgksvllsvggekavavnds--saqrfadst 107  
QY 160 --LVDELGFDGIDVYEPNGSEFDGLNDEKKADEFQVYTKLREVMCDKRLISISQSSNGA 217  
DB 108 yalmeeygfgyvdldle-----ngln-----stymtealklhekagdgvlvmagptc--- 155  
QY 218 LSCIGFNDPKKICMDEAPYNSKYFNKPDVKKELLRA-----AQMASAGCAIYLMNN 269  
DB 156 -----ldmspen-eyfktaivtkdflavmgyngsgmslgcdggyayagt 201  
QY 270 LKDMIDMVFVQTFN 283  
DB 202 vdflltalaciqten 215

RESULT 4  
AA63552 standard; Protein: 298 AA.  
XX  
XX AAG63552;  
XX  
XX 15-OCF-2001 (first entry)  
XX  
XX Amino acid sequence of a chitinobiosidase polypeptide.  
DE  
XX Chitinobiosidase; chitinolytic enzyme; early flowering.  
XX  
XX Streptomyces albidoflavus.

XX WO200146387-A1.  
PN  
XX 28-JUN-2001.  
XX  
XX 22-DEC-2000; 2000WO-US35238.  
XX  
XX 23-DEC-1999; 99US-0172003.  
XX  
XX (COR ) CORNELL RES FOUND INC.  
XX  
XX Broadway RM, Gongora CE;  
PI  
XX WPI: 2001-496661/54.  
DR N-PSDB; AAH74537.  
XX  
XX Promoting early flowering, yield in plants and reducing plant size,  
PT involves growing transgenic plant or plant seed transformed with DNA  
PT molecule encoding chitinolytic enzyme, or applying the enzyme to the  
PT plant  
XX  
XX Claim 7; Page 10-11; 87pp; English.  
XX  
XX The present sequence represents a chitinobiosidase polypeptide. The  
CC polypeptide has a molecular mass of 34 kD and an isoelectric point  
CC of less than 3. The polypeptide is a chitinolytic enzyme. The  
CC chitinolytic polynucleotide is used to produce transgenic early  
CC flowering plants. The chitinolytic polynucleotides and polypeptides  
CC are useful for promoting early flowering in plants, promoting yield  
CC from plants and reducing plant size.  
XX  
XX Sequence 298 AA;  
SQ  
Query Match 6.8%; Score 135.5; DB 22; Length 298;  
Best Local Similarity 21.7%; Pred. No. 0.0002;  
Matches 55; Conservative 50; Mismatches 76; Indels 73; Gaps 12;  
QY 55 IQGYPSWVSYNH-----NLKDLNPNLVNVMSPFAKMDLSYDSISIVGSPLLFKSLGL 109  
DB 10 vlgv---wqnfngatvgtladvdpaydlavsfadatanagel-----tftl 54  
QY 110 EYIGLNEFYNDAMNLRKRPDI-----MLSLGGETYHPSPSDSLANVEKIAN-- 159  
DB 55 dsvglggytde-----qfradlaakqadgksvllsvggekavavnds--saqrfadst 107  
QY 160 --LVDELGFDGIDVYEPNGSEFDGLNDEKKADEFQVYTKLREVMCDKRLISISQSSNGA 217  
DB 108 yalmeeygfgyvdldle-----ngln-----stymtealklhekagdgvlvmagptc--- 155  
QY 218 LSCIGFNDPKKICMDEAPYNSKYFNKPDVKKELLRA-----AQMASAGCAIYLMNN 269  
DB 156 -----ldmspen-eyfktaivtkdflavmgyngsgmslgcdggyayagt 201  
QY 270 LKDMIDMVFVQTFN 283  
DB 202 vdflltalaciqten 215

RESULT 5  
AA52533 standard; Protein: 509 AA.  
XX  
XX AA52533;  
XX  
XX 22-FEB-2000 (first entry)  
XX  
XX D. pteronyssius 98 kD mite allergen protein (map) PDerp98-509.  
DE  
XX Mite allergen protein; map; high molecular weight; HMW-map; allergy;  
XX house dust mite; IGE; immunoglobulin E; allergen; mapB;  
XX hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;  
XX canine; veterinary; antibody; vaccine; immunisation.

```

XX OS Dermaphagoides pteronyssius.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT Protein /note="Signal peptide"
XX FT /note="Mature Pderp98-509"
XX PN WO9954349-A2.
XX PD 28-OCT-1999.
XX PF 16-APR-1999; 99MO-US08524.
XX PR 17-APR-1998; 98US-0062013.
XX PR 13-MAY-1998; 98US-0085295.
XX PR 02-SEP-1998; 98US-0098909.
XX PA (HESK-) HESKA CORP.
XX PI McCall CA, Hunter SW, Weber ER;
XX DR WPI: 2000-052700/04.
XX DR N-PSDB: AA238585, AA238586, AA238587, AA238588.
XX PT Novel high molecular weight Dermatophagoides nucleic acid polypeptides
XX PT used to modify an animals' hypersensitivity to mite allergens
XX PS Claim 3; Page 134-136; 154pp: English.
XX CC This sequence represents Dermatophagoides pteronyssius mite allergen
XX CC protein (map) Pderp98-509. Pderp98-509 has a molecular weight of 98 kD,
XX CC comprising 509 amino acids, and has a high degree of homology with the
XX CC D. farinae 98 kD allergen, mapB (AAV52523). Nucleic acid molecules
XX CC encoding Pderp98-509 were isolated from a D. pteronyssius cDNA library
XX CC by hybridisation with a probe encoding the D. farinae high molecular
XX CC weight map (Hmw-map) composition. Mite allergenic proteins and
XX CC peptides, and nucleic acids encoding them, may be used in therapeutic
XX CC compositions to modify an animal's hypersensitivity reaction to mite
XX CC allergens. Animals that may be treated include mammals and birds,
XX CC especially felines, canines, equines, humans, other pets, and work or
XX CC domestic animals. The proteins or fragments may also be used to
XX CC diagnose allergies via a skin test. The proteins and peptides can also
XX CC be used to raise antibodies, which have a variety of potential
XX CC uses. For example, they can be used as vaccines to passively immunise
XX CC animals against dust mite hypersensitivity, as positive controls in
XX CC test kits and as tools to recover desired dust mite allergens from a
XX CC mixture of proteins.
XX SQ Sequence 509 AA:

```

Query Match 6.8%; Score 134.5; DB 21; Length 509;  
 Best Local Similarity 22.1%; Pred. No. 0.00054;  
 Matches 83; Conservative 51; Mismatches 128; Indels 113; Gaps 20;

```

OY 5 VVYSFLVLCCLLSTVYSVIEGHRARPEGRNPRELITKTESGKGIIGYPPSW 64
DB 1 mktfalciwajgmnaatkrdh---nysknpmriv-----cyvgtwsv 44
OY 65 YN---HNKLDLPNLNVVHM--SFAKMDLSYDSIESIVGSPLLFKSLIGLEYGLN 118
DB 45 yhvvdqytltdldp-fkcthlmygfkideykyltg--vtdpfdqdhmswekhgy-erf 100
OY 119 NQAMNRRKARPDILIMLSLGG-----ETVHPSFDSALNA-----VEKINLVDELG 167
DB 101 n---nlrlknpeitmtisjgwyegsekys-----dmaapntyrgqfvgsvldifgeykfd 153
OY 168 GIDVDYEPNGSGFDLNDKEKADFFVOVYTKLREYMCDDKLISISQSSNGALSCIGFNDPK 227
DB 154 gldldweygsrslg-npkldkgnyltlvrelke-----afepgyg---- 192

```

```

OY 228 KICMDDEAPYNSKYFNKPDVKELLRAAQAASAG-----GAIYLMNNLKMDIDWVYOTFN 283
DB 193 -----lltaa--vspgkdkidvayelkelnglfdmnmvntld 227
OY 284 YTNSTDSFV-----MKELYSYAYYGKKY--DVYIIMGFP-----LMPPSPFPN 325
DB 228 yhgwenvfghmaplkykpdetdelhcyfnvymhyllmgatrdklmgvpyfgyraws 287
OY 326 PNDKMLVKSIGDFVK 340
DB 288 ldrskvk-lgdpak 301

```

RESULT 6  
 AAU00962  
 ID AAU00962 standard; Protein: 399 AA.  
 AC AAU00962;  
 DT 12-SEP-2001 (first entry)  
 DE F. venenatum endochitinase polypeptide.  
 XX Glucanase; endochitinase; exochitinase; cell-wall degradation; fungus;  
 KW transgenic plant; plant pathogen; bacteria; seafood waste; shell;  
 KW chitin; chemical modification; glucan.  
 XX Fusarium venenatum.  
 XX MO200116353-A1.  
 XX PD 08-MAR-2001.  
 XX PF 30-AUG-2000; 2000MO-US23802.  
 XX PR 30-AUG-1999; 99US-0151582.  
 XX PR 11-AUG-2000; 2000US-0224946.  
 XX PR 28-AUG-2000; 2000US-0649747.  
 XX (NOVO ) NOVO NORDISK BIOTECH INC.  
 XX (USDA ) US SEC OF AGRIC.  
 XX Okubara PA, Blechl AE, Hohn TM, Berka RM;  
 XX WPI: 2001-218524/22.  
 XX N-PSDB: AAS00796, AAS00800.  
 XX Fusarium nucleic acids encoding polypeptides having glucanase,  
 PT endochitinase or exochitinase activity, useful for producing transgenic  
 PT plants which are resistant to plant pathogens, particularly Fusarium  
 PT species -  
 XX Claim 13; Page 167-168; 216pp; English.  
 XX The sequence represents the Fusarium fungal enzyme, endochitinase.  
 XX Glucanase, endochitinase and exochitinase are polypeptides with cell-wall  
 XX degrading activity, derived from Fusarium fungal genes. The associated  
 XX nucleic acids can be used to produce transgenic plants which are  
 XX resistant to plant pathogens, particularly Fusarium species. They can  
 XX also be used to isolate homologous genes from fungi to obtain genes which  
 XX protect host cells, including fungi, bacteria and plants against related  
 XX fungal pathogens. The polypeptides, especially chitinases and glucanases,  
 XX are useful for degrading seafood waste, such as shells that contain  
 XX chitin, or for chemical modification of chitin or glucan.  
 XX SQ Sequence 399 AA:

Query Match 6.8%; Score 130; DB 22; Length 399;  
 Best Local Similarity 22.4%; Pred. No. 0.001;  
 Matches 73; Conservative 51; Mismatches 112; Indels 90; Gaps 16;

```

OY 58 YVPSWVSY--NHNKLDLPNLNVVHMSFAKMDLSYDSIESIVGSPLLFKSLIGLEYGLN 115

```



```

Db      13 yfwnaiyatkhhpqdl-pventhilly-----sfanirsdsgvnhltdswadtdimdg 66
QY      116 EYENDA-----MNLKRRAR-PDITMLSLGGETYH-----PSSFSDSALNAVER-IA 158
Db      67 dswndvgtlnylgcmkqnlklkrnmlklsigwtfssnfkgpatpggrdtfakscv 126
QY      159 NLVDELGFDCIDVDYE-PNGSFGDLNDKEADFFVOYVTKLREYMCDDKLISQSSNGA 217
Db      127 dlknlgfdgididweypr-----qdaearnlyvellgavrhnm-dayaqtlisq----- 173
QY      218 LSGIGFNDPKKICMDDEAPYNSKYFNKPDKKELLRQAQNASGAIYLMNNLKMDIMV 277
Db      174 -----pyhfeltvaecpagaunfqlklr-----gmduqyldfw 205
QY      278 FVOYFNTNSTDSTV--MKELDYSAVYGGKRYVILMGFTLMEPSYFPNPKMLKSI 335
Db      206 nlmaydyagwdqtaghqnalypr-----hdnpr-----stpfacts-----aal 244
QY      336 GDVVKTEENKLNKRADGFGMLSSSDN 361
Db      245 dlyvrsqvmpekivlgmplygrafen 270

```

## RESULT 7

```

AAW29178
ID      AAW29178 standard; Protein; 561 AA.
AC      AAW29178;
XX
XX
DT      06-MAR-1998 (first entry)
DE
XX
XX      S. marcescens chitinase A.
KM      Serratia marcescens; chitinase A; transgenic plant; cold resistant;
KW      reducing sugar; increased level.
XX
OS      Serratia marcescens.
XX
PN      US5633450-A.
XX
PD      27-MAY-1997.
XX
PE      01-DEC-1995; 9505-0566347.
XX
PR      19-DEC-1994; 94US-0358901.
PR      26-MAR-1984; 84US-0593691.
PR      18-JUL-1986; 86US-0888033.
PR      09-JUL-1990; 90US-0550253.
PR      14-AUG-1992; 92US-0930970.
PR      01-DEC-1995; 9505-0566347.
XX
XX      (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
XX      Jones JDC, Suslow TV;
XX
XX      WPI; 1997-297423/27.
XX      DR      N-PSDB; AAT86836.
XX
XX      Transgenic plants expressing chitinase - with increased cold
XX      PT      resistance or sugar content
XX
XX      Example 3; Columns 25-30; 25pp; English.
XX
XX      This sequence represents the Serratia marcescens chitinase A.
XX      CC      information from the DNA coding it (AAT86836) was used in the creation of
XX      CC      constructs which led to the expression of the bacterial chitinase gene in
XX      CC      plant cells. Plants which have been transformed with chitinase-encoding
XX      CC      DNA so that they express active chitinase are resistant to cold damage or
XX      CC      have enhanced reducing sugar levels. The plants, especially tomato
XX      CC      plants, have improved frost resistance or produce sweeter fruits.
XX
SQ      Sequence 561 AA.

```

Query Match 6.5%; Score 128.5; DB 18; Length 561;  
 Best Local Similarity 23.2%; Pred. No. 0.0023;  
 Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

```

QY      2 NFTV-----KISFLVYICLLC-----CLSTYVSVT-----EGHARARGESRKNP-REIITKTFK 48
Db      99 nfkvnkgqrgymqv--alcnadgctasdateivvadtdgrhlpp---lkepllleknkpyk 153
QY      49 ESGKGIIOGYVPSWVSYNHNK-LDNPMLNVHMSFAKMDIS-----YDSIESTIGS-PL 101
Db      154 gnsqkvvgysyivewgyvgrnftvdkipeqnlthlygfpicgngndstikelegstqa 213
QY      102 LFKSLIGLEVIGLNEY-----FNDANNLRKARPDITMLSLG 138
Db      214 lqrcqgredtklshdpfaalqkqgvtawddpykgnfglnalqahpdlkllp9ig 273
QY      139 GETYHPSSF-----DSALNAVEKIANLVDELGFDCIDVDYE---PNGSFGDLNDKEK 187
Db      274 gwltisdpffimgdkvkdrifgsvkef--lqtwkffidgvdidwefpgqkganpnlgsppd 331
QY      188 ADFEVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMDDEAPYNSKYFNKPDPV 247
Db      332 getyvlmkeiraml--dql-----ssetgrky----- 357
QY      248 KKELLRAAQAQNASGAIYLMNNLKMDIMVFOYFNT--NSTDSTVMEKYDYSAAYG 304
Db      358 --eltsa---lsag-----kdkldkv---aynvagmsmdhlflm---sydfyg 394

```

## RESULT 8

```

AAR64823
ID      AAR64823 standard; Protein; 561 AA.
XX
XX      AAR64823;
XX
XX
DT      23-JUL-1995 (first entry)
DE
XX
XX      Chitinase A.
KM      Chitinase A; protein secretion; biological control agent;
KW      transgenic plant; chitin; plant pathogen; nematode; fungus;
XX      crop protection; PCHIT1251.
XX
OS      Serratia marcescens QMB1466.
XX
FH      Key      Location/Qualifiers
FT      Peptide 1..23 /label= sig_peptide
XX
XX      US5374540-A.
XX
XX      20-DEC-1994.
XX
XX      26-MAR-1984; 84US-0593691.
XX
XX      26-MAR-1984; 84US-0593691.
XX      PR      18-JUL-1986; 86US-0888033.
XX      PR      09-JUL-1990; 90US-0550253.
XX
XX      (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
XX      Jones JDC, Suslow T;
XX
XX      WPI; 1995-035648/05.
XX      DR      N-PSDB; AAO76290.
XX
XX      Plant cell conty. signal sequence of chitinase A gene - for
XX      PT      secretion of foreign polypeptide, esp. chitinase A for protection
XX      PT      against pathogenic fungi and nematodes.
XX
PS      Disclosure; Fig. 1A-1C; 20pp; English.

```

XX Serratia marcescens QMB1466 genomic DNA was partially digested,  
CC ligated into vector pLARI, packaged into lambda phage particles,  
CC and used to transfect *Escherichia coli* DH1 (ATCC 33849) cells.  
CC Plating on chitin medium was used to screen for chitinase activity.  
CC Clone C12 (ATCC 67152) was isolated. Its sequence is given in  
CC AA076290, and the deduced protein sequence in AAR64823.

XX  
SQ Sequence 561 AA:

Query Match 6.4%; Score 127.5; DB 16; Length 561;  
Best Local Similarity 23.2%; Pred. No. 0.0028;  
Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

QY 2 NFTV---KSYFLYICLLC---CLSTYVSVI---EGHRAPGESRKNP-REIITFK 48  
DB 99 ntkvknkgrygmqv--alcnaadgctasdateivvadtdgshlpp---lkeplleknpkyk 153  
QY 49 ESGKGIIQGYPSWVSYNHNLK-DLNPNLNVHMSFAKMDLS-----YDSIESIVGS-PL 101  
DB 154 qnsqkvvgysfvegwgygnftvdklpagqlthlygfipicgngindslkelegsfga 213  
QY 102 LFKSLGLEIYIGLNEY-----FNDAMNLRKARPDIIMLSLG 138  
DB 214 lqrsccgredfkisldpfcaalqkaqgyrawddpykgnfgqlmalqahpdlkllpsig 273  
QY 139 GETYHPSSP-----DSALNVEKIANLVDELFGDIDVDE---PNCSPFGCLNDKER 187  
DB 274 gwtlsdppffimgdkvkrdfrvsgvkcf--lgtwkffdgvdidwefpgykganpnlgspqd 331  
QY 188 ADFEVOYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMDDEAPNRSKYFNKPDV 247  
DB 332 getyvllmkelraml--dqj-----saetgryk----- 357  
QY 248 KKEILRAQMASAGAIYLMNNIKMDIMVYOTFNVT-NSTDSTYWKELYSYAYYG 304  
DB 358 --eltsa--lsag-----kdkidkv--aynvaqnsmdhlfilm---sydtyg 394

#### RESULT 9

ID AAM09402 standard; Protein; 561 AA.

AC AAM09402;

DT 19-FEB-1997 (first entry)

DE Serratia marcescens Chitinase A.

XX Chitinase; ChIA; chitin degradation; transgenic plant; tomato;

KW cold resistance; tolerance; frost damage; reducing sugar content;  
KW increase; ATCC 9901; ds.

XX Serratia marcescens QMB 1466.

OS Serratia marcescens QMB 1466.

XX Key Location/Qualifiers

FT Peptide 1..23 /label= signal\_peptide

FT /note= "predicted cleavage site is between residues

FT Protein 24..561 /label= Chitinase\_A

FT /note= "predicted mature sequence"

XX US554521-A.

XX 10-SEP-1996.

XX 26-MAR-1984; 84US-0593691.

XX 19-DEC-1994; 94US-0358901.

XX 26-MAR-1984; 84US-0593691.

PR 18-JUL-1986; 86US-0888033.  
PR 09-JUL-1990; 90US-0550253.  
PR 14-AUG-1992; 92US-0930970.

XX (DNAP ) DNA PLANT TECHNOLOGY CORP.

XX Jones JDC, Suslow TV;

XX WPI: 1996-424655/42.

DR N-PSDB; AAT49479.

PT Recombinant plant with increased resistance to cold - and increased  
PT reducing sugar content, contains DNA encoding chitinase, partic. to  
PT protect tomatoes against frost damage.

XX Example 3; Fig 1A-C; 25pp; English.

CC The present sequence is the Chitinase A enzyme encoded by the 2.3 kb  
CC insert of plasmid pCHIT1251 which was isolated from genomic DNA of  
CC Serratia marcescens QMB 1466 (ATCC 9901). The chIA gene can  
CC be used for transforming plants, esp. tomatoes, to give increased  
CC resistance to frost damage and to produce sweeter fruit.

XX Sequence 561 AA;

Query Match 6.4%; Score 127.5; DB 17; Length 561;  
Best Local Similarity 23.2%; Pred. No. 0.0028;  
Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

QY 2 NFTV---KSYFLYICLLC---CLSTYVSVI---EGHRAPGESRKNP-REIITFK 48  
DB 99 ntkvknkgrygmqv--alcnaadgctasdateivvadtdgshlpp---lkeplleknpkyk 153  
QY 49 ESGKGIIQGYPSWVSYNHNLK-DLNPNLNVHMSFAKMDLS-----YDSIESIVGS-PL 101  
DB 154 qnsqkvvgysfvegwgygnftvdklpagqlthlygfipicgngindslkelegsfga 213  
QY 102 LFKSLGLEIYIGLNEY-----FNDAMNLRKARPDIIMLSLG 138  
DB 214 lqrsccgredfkisldpfcaalqkaqgyrawddpykgnfgqlmalqahpdlkllpsig 273  
QY 139 GETYHPSSP-----DSALNVEKIANLVDELFGDIDVDE---PNCSPFGCLNDKER 187  
DB 274 gwtlsdppffimgdkvkrdfrvsgvkcf--lgtwkffdgvdidwefpgykganpnlgspqd 331  
QY 188 ADFEVOYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMDDEAPNRSKYFNKPDV 247  
DB 332 getyvllmkelraml--dqj-----saetgryk----- 357  
QY 248 KKEILRAQMASAGAIYLMNNIKMDIMVYOTFNVT-NSTDSTYWKELYSYAYYG 304  
DB 358 --eltsa--lsag-----kdkidkv--aynvaqnsmdhlfilm---sydtyg 394

#### RESULT 10

ID AAM60179 standard; Protein; 561 AA.

AC AAM60179;

DT 03-SEP-1998 (first entry)

DE Serratia marcescens chitinase A sequence.

XX Chitinase A; pathogen; plant protection; Serratia marcescens; E. coli;

KW Pseudomonas; Erwinia; frost damage resistance; plant pathogen; fungus;

XX Serratia marcescens.

XX Key Location/Qualifiers

FT Peptide 1..23

```

FT      /note= "putative signal peptide"
FT      23..24
FT      Cleaveage-site /note= "predicted cleavage site"
FT      Protein 24..561
FT      /note= "mature protein"
XX
XX      US5776448-A.
XX
XX      07-JUL-1998.
XX
XX      01-AUG-1996; 96US-0693835.
XX
XX      19-DEC-1994; 94US-0358901.
XX      26-MAR-1984; 84US-0593691.
XX      18-JUL-1986; 86US-0888033.
XX      09-JUL-1990; 90US-0550253.
XX      14-AUG-1992; 92US-0930970.
XX      01-AUG-1996; 96US-0693835.
XX
XX      (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
XX      Jones JDC, Suslow TV;
XX
XX      WPI: 1998-397941/34.
XX      N-PSDB; AAV35601.
XX
XX      Use of heterologous chitinase nucleic acid - for transforming
XX      bacteria or plants, for increasing plant resistance to plant
XX      pathogens, increase chilling resistance and increasing sweetness.
XX
XX      Example 3; Fig 1A-C; 25pp; English.
XX
XX      This represents a Serratia marcescens chitinase A. The invention
XX      provides a method for the production of a bacterial cell capable of
XX      producing chitinase which comprises transforming the bacterial cells
XX      with a nucleic acid encoding the chitinase, the nucleic acid being
XX      isolated from a heterologous source, and whereby the bacterial cell
XX      is capable of expressing the nucleic acid. The bacterial cell is selected
XX      from Pseudomonas sp., E. coli or Erwinia sp. The products can be used to
XX      enhance plant growth by biological control of plant pathogens such as
XX      fungi, nematodes, insects and disease agents. Plants transformed with
XX      the chitinase DNA can also have resistance to frost (freezing) damage
XX      or chilling damage, increased levels of reducing sugars and sweetness
XX      in fruits or plants and enhanced post-harvest storage life. The
XX      products can also be used for the production of chitinase for use as an
XX      antibiotic.
XX
XX      Sequence 561 AA:
XX
XX      Query Match 6.4%; Score 127.5; DB 19; Length 561;
XX      Best Local Similarity 23.2%; Pred. No. 0.0028;
XX      Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

```

```

Oy      248 KELLRAQMASGALYIMNLIKMDIDVQTEFNT-NSDSTVMEKELYDVAAYG 304
Oy      358 --eltsa---lsag-----kdkidkv---aynvagmsdhflfm----sydfyg 394
Db
RESULT 11
AAV52523
ID      AAV52523 standard; Protein; 555 AA.
XX
XX      AAV52523;
XX
XX      22-FEB-2000 (first entry)
XX
XX      House dust mite (D. farinae) mite allergen protein (map) Pderf98-555.
XX
XX      Mite allergen protein; map; high molecular weight; HMW-map; allergy;
XX      house dust mite; IgE; immunoglobulin E; allergen; mapA; mapB;
XX      hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
XX      canine; veterinary; antibody; vaccine; immunisation.
XX
XX      Dermatophagoides farinae.
XX
XX      Key Location/Qualifiers
XX      Peptide 1..19
XX      Protein /note= "Signal peptide"
XX      FT 20..555
XX      FT /note= "Mature Pderf98-555"
XX
XX      WO9954349-A2.
XX
XX      28-OCT-1999.
XX
XX      16-APR-1999; 99WO-US08524.
XX
XX      17-APR-1998; 98US-0062013.
XX      13-MAY-1998; 98US-0085295.
XX      02-SEP-1998; 98US-0098909.
XX
XX      (HESK-) HESKA CORP.
XX
XX      McCall CA, Hunter SM, Weber ER;
XX
XX      WPI: 2000-052700/04.
XX      N-PSDB; AAZ38575, AAZ38576, AAZ38577, AAZ38578.
XX
XX      Novel high molecular weight Dermatophagoides nucleic acid polypeptides
XX      used to modify an animals' hypersensitivity to mite allergens -
XX
XX      Claim 3; Page 111-113; 154pp; English.
XX
XX      This sequence represents Dermatophagoides farinae mite allergen protein
XX      (map) Pderf98-555. Pderf98-555 has a molecular weight of 98 kD,
XX      comprising 555 amino acids, and is a component of the Dermatophagoides
XX      farinae high molecular weight mite allergen protein (HMW-map)
XX      composition. The HMW-map composition was isolated from a D. farinae
XX      homogenate by gel filtration, with each fraction being analysed for
XX      the presence of proteins that bound to IgE present in mite-allergic
XX      dog antisera. Mite allergenic proteins and peptides, and nucleic acids
XX      encoding them, may be used in therapeutic compositions to modify an
XX      animal's hypersensitivity reaction to mite allergens. Animals that
XX      may be treated include mammals and birds, especially felines,
XX      canines, equines, humans, other pets, and work or domestic
XX      animals. The proteins or fragments may also be used to diagnose
XX      allergies via a skin test. The proteins and peptides can also
XX      be used to raise antibodies, which have a variety of potential
XX      uses. For example, they can be used as vaccines to passively immunise
XX      animals against dust mite hypersensitivity, as positive controls in
XX      test kits and as tools to recover desired dust mite allergens from a
XX      mixture of proteins.
XX
XX      Sequence 555 AA:
XX
XX

```

Query Match 6.4%; Score 126; DB 21; Length 555;  
 Best Local Similarity 21.2%; Pred. No. 0.0039;  
 Matches 95; Conservative 57; Mismatches 160; Indels 136; Gaps 25;

```

OY 5 VKISFLVICLLCLLSTVSYIEGHRARPGESRKNPREITTFEESGKGIIGYPSNVS 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   1 mktiyalislsmactglmaslkrth-----ndyskpmrtriv-----cygwtsv 44
DB
OY 65 YN-----HNLEDLNPNLNVVHM--SFAMKDLSDYSIESIVGSPLLFKSLIGLEYIGLN 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   45 yhkvdpyttiedip-fctchlmvgfakideyktyiq--vfipygdannswekrgy-erf 100
OY 119 NDAMNLRKARPDITIMLSLGG-----ETYPSSFDLSALNA-----VEKIANLVDELGFD 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   101 n--nrlrknpelttmislglgwygsekys-----dmanptryqglsgvldfdeykfd 153
DB
OY 168 GIDVDEPNNSFDG-----LND-----KEKAD--FFVQY 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   154 gllddeyppsrjgnphkdkqnylalvrelkdafebhylldaavspgkdkidraydike 213
OY 195 VTKLREVMCDKLISISQSSNGALSCIGFNDP--KKICMDDE-----APYNSKYF--NKP 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   214 lnlldfwm-----nvmtydygngwenfygnaplykrpdetelhtyfnvnytmbyhlyng 269
OY 246 DVKELIRAAQMASAGAIYLMNML-----KDMIDVFOV-----TFNYTSTDSVTNKE 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   270 atrdklvmgvpfygrawledrsklikgdpakmsppglsgvslslelqclfgkqe 329
OY 296 LVDSY--AYGKKYDYVIMGFILMFSPFPNPKMLVKSIGDFVKTENKL-----NKKAD 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   330 whqydeyynapygy-----ndkltw-gyddlaaslscklaflkelgvs 371
DB
OY 351 GFGLWSTLSD-----NAAHN 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   372 gymvwslenddfkghcgkpnllnkvh 399
DB

```

## RESULT 12

AAR63441  
 ID AAR63441 standard; Protein; 424 AA.

XX AAR63441;

DT 06-JUL-1995 (first entry)

DE Trichoderma harzianum P1 endochitinase.

KW Trichoderma harzianum P1 endochitinase: inhibition;

KW chitin containing fungi; herbivorous insects.

OS Trichoderma harzianum.

PN WO9424288-A.

PD 27-OCT-1994.

PF 10-FEB-1994; 94WO-US01198.

PR 14-APR-1993; 93US-0045269.

PA (CORR ) CORNELL RES FOUND INC.

PI Harman GE, Hayes CK, Lorito M, Tronsmo A;

DR WPI: 1994-341867/42.

DR N-PSDB; AAQ78144.

PT Trichoderma harzianum P1 endo:chitinase and chitinase genes -  
 useful in biological control of chitin-contg. fungi and insects  
 (Eng).

PS Example 8; Page 45; 62pp: English.

XX AAQ78144 encodes AAR63441 the Trichoderma harzianum P1 endochitinase,  
 CC which can be used to inhibit chitin containing fungi and chitin  
 CC containing herbivorous insects, especially from the genera below  
 CC Fusarium, Gliocladium, Rhizoctonia, Trichoderma, Uncinula, Ustilage,  
 CC Erysiphe, Botrytis, Sclerotium, Sclerotium and Alternaria.  
 SO Sequence 424 AA;

Query Match 6.3%; Score 124.5; DB 15; Length 424;  
 Best Local Similarity 21.0%; Pred. No. 0.0036;  
 Matches 65; Conservative 51; Mismatches 93; Indels 101; Gaps 14;

```

OY 58 YPSPVSNVHNLKDLN--PNLNVVHMSFAKMDL--SYDSIESIVGSPLLFKSLIGLEYIGLN 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   44 yftmglygrnlfqpnllvasdltlhyisfmlleaygrvs-----gdayadyq 91
DB
OY 116 EYFND-----AMNLRKARPDITIMLSLGGCTTYPSSFDLSAL-----N 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   92 khyddsdvndvgnnagvcvklqfklkknngllkwmllsgwtw--stnfpaastdanrkn 150
DB
OY 153 AVEKIANLVDELGFDGIDVDEPNNSFDGLNDKEKADFEVQYVTKLREVMCDKLISISQ 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   151 faktaitfmkdwgfydgldvdr-----eypad-----tq 179
DB
OY 213 SSNGALSCIGFNDPKKICMDDEAPYNSKYFKNPDYKELIRAAQMASAGAIYLMNMLKD 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   180 atnmvlll-----ketrsglda-yaagy--apgyhflslsapapehysflmsdlgq 230
OY 273 MIDVFOVTFNYTSTDSVTNKELYDSYAYGKKYDYVIMGFILMFSPFPNPNDR--M 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   221 vldvynlmaydags-----wssys-----ghdanlfaupnpspyn 269
DB
OY 331 LVKSIGDFVK 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   270 tdgaikgyik 279
DB

```

## RESULT 13

AA52535  
 ID AA52535 standard; Protein; 490 AA.

XX AA52535;

DT 22-FEB-2000 (first entry)

DE D. pteronyssius 98 kD mite allergen protein (map) pDerp98-490.

KW Mite allergen protein; map; high molecular weight; HMW-map; allergy;

KW house dust mite; IGE; immunoglobulin E; allergen; mapB;

KW hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;

KW canine; veterinary; antibody; vaccine; immunisation.

OS Dermatophagoides pteronyssius.

PN WO9954349-A2.

PD 28-OCT-1999.

PF 16-APR-1999; 99WO-US08524.

PR 17-APR-1998; 98US-0062013.

PR 13-MAY-1998; 98US-0085295.

PR 02-SEP-1998; 98US-0098909.

PA (HESK-) HESKA CORP.

PS











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 08:57:07 ; Search time 12.57 Seconds

(without alignments)  
676,710 Million cell updates/sec

Title: US-09-579-383-3

Perfect score: 1982

Sequence: 1 MNFTVYKSLVLCILCLLS.....SDNAHNEQLATFVFSLSH 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*

6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	135.5	6.8	298	3	US-09-025-691-1
2	127.5	6.4	561	1	US-08-358-901-2
3	127.5	6.4	561	1	US-08-566-347-2
4	127.5	6.4	561	1	US-08-693-835-2
5	124.5	6.3	424	5	PCT-US94-01198-2
6	120.5	6.1	424	1	US-08-045-269C-2
7	120.5	6.1	424	3	US-08-371-680-2
8	117.5	5.9	866	1	US-08-386-727-8
9	117.5	5.9	866	2	US-08-600-452A-8
10	111	5.6	461	2	US-08-392-625-24
11	111	5.6	461	2	US-08-466-961A-24
12	111	5.6	461	2	US-08-645-193B-26
13	110.5	5.6	376	3	US-09-025-691-3
14	108.5	5.5	1453	1	US-08-308-872B-6
15	107	5.4	554	2	US-08-524-051-2
16	107	5.4	554	2	US-09-052-778-16
17	102	5.1	327	2	US-08-926-724-3
18	102	5.1	327	3	US-08-904-032-3
19	101.5	5.1	1046	1	US-08-386-727-2
20	101.5	5.1	1046	2	US-08-600-452A-2
21	100.5	5.1	423	1	US-07-939-501A-10
22	100.5	5.1	423	1	US-07-939-501A-12
23	100	5.0	440	3	US-09-052-778-13
24	99.5	5.0	389	1	US-07-939-501A-1
25	99.5	5.0	389	4	US-08-448-398-7
26	99.5	5.0	1452	3	US-08-331-625A-2
27	99.5	5.0	1452	5	PCT-US93-04384-18

28	99.5	5.0	1452	5	PCT-US93-04692-2	Sequence 2, Appli
29	98	4.9	128	3	US-08-331-625A-51	Sequence 51, Appl
30	98	4.9	250	3	US-08-331-625A-57	Sequence 57, Appl
31	98	4.9	327	2	US-08-926-724-1	Sequence 1, Appli
32	98	4.9	362	3	US-08-331-625A-53	Sequence 53, Appl
33	98	4.9	1443	1	US-08-308-872B-2	Sequence 2, Appli
34	98	4.9	1454	4	US-08-392-459-32	Sequence 32, Appl
35	98	4.9	1454	4	PCT-US91-08525-32	Sequence 32, Appl
36	98	4.9	1454	5	PCT-US93-04384-12	Sequence 43, Appl
37	98	4.9	1454	5	PCT-US93-04384-43	Sequence 46, Appl
38	98	4.9	1454	5	PCT-US93-04384-46	Sequence 48, Appl
39	98	4.9	1454	5	PCT-US93-04384-48	Sequence 1, Appli
40	97.5	4.9	720	2	US-08-840-236-1	Sequence 1, Appli
41	97.5	4.9	720	2	US-08-505-448A-1	Sequence 1, Appli
42	97.5	4.9	1346	1	US-08-471-053-23	Sequence 23, Appl
43	97.5	4.9	1346	2	US-08-471-044-23	Sequence 23, Appl
44	97.5	4.9	1346	2	US-08-463-483A-23	Sequence 23, Appl
45	97.5	4.9	1346	2	US-08-471-046A-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-09-025-691-1
Sequence 1, Application US/09025691
Patent No. 6069299
GENERAL INFORMATION:
APPLICANT: Broadway, Roxanne M.
TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH
TITLE OF INVENTION: CHITINOLYTIC ENZYMES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025, 691
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/20120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-691-1

Query Match 6.8% Score 135.5; DB 3; Length 298;
Best Local Similarity 21.7% Pred. No. 2.5e-05;
Matches 55; Conservative 50; Mismatches 76; Indels 73; Gaps 12;

Oy 55 IOGYPPSVNHN-----NLKDLNPNLVNMFKMLSDYSISYIGSPILFSLGL 109
Db 10 VTGT---WQNNKATVOTLADVPDAVDIIVAFADATAMAGET-----FTTL 54
```

```

QY 110 EYGLNKEFNAMNRRKRRPOLI-----MLSLGEGYTHPSFSALNAVEKIAN-- 159
Db 55 DSVGGGTTD-----QFRADLANQADKQSVIISVGGKCAVANVDA--SAQREFADST 107
QY 160 --LVDELGFDSIDVDFEFGSFGDGNKEKADFVQVYVTKLREYMCDDKLISQSNGA 217
Db 108 YALMEYEFDCGVIDILE-----NGIN----STYMTALTKLHEKAGGGLVLTMAPQ--- 155
QY 218 LSCIGFNDPKKICMDEAPYKSYFNKRPDYKELLRA-----AOMASAGAIYLMNN 269
Db 156 -----IDMOSPEN-EYFKTALVTKDCLTAVNMQYNSGSMLCGDGVVNOGT 201
QY 270 LKQMDMWFVQTFN 283
Db 202 VDFLTALACIDEN 215

```

```

1      RESULT      2
2      US-08-358-901-2
3      : Sequence 2, Application US/083586901
4      : Patent No. 5554521
5      : GENERAL INFORMATION:
6      : APPLICANT: Suslow, Trevor V.
7      : APPLICANT: Jones, Jonathan D.G.
8      : TITLE OF INVENTION: No. 5554521el Chitinase-Producing Plants
9      : NUMBER OF SEQUENCES: 10
10     : CORRESPONDENCE ADDRESS:
11     : ADDRESSEE: Townsend and Townsend Khourie and Crew
12     : STREET: One Market Plaza, Stewart Tower, Suite 2000
13     : CITY: San Francisco
14     : STATE: California
15     : COUNTRY: USA
16     : ZIP: 94105
17     :
18     : COMPUTER READABLE FORM:
19     : MEDIUM TYPE: Floppy disk
20     : COMPUTER: IBM PC compatible
21     : OPERATING SYSTEM: PC-DOS/MS-DOS
22     : SOFTWARE: Patentin Release #1.0, Version #1.25
23     : CURRENT APPLICATION DATA:
24     : APPLICATION NUMBER: US/08/358,901
25     : FILING DATE: 19-DEC-1994
26     : CLASSIFICATION: 435
27     : PRIOR APPLICATION DATA:
28     : APPLICATION NUMBER: US 07/550,253
29     : FILING DATE: 09-JUL-1990
30     : PRIOR APPLICATION DATA:
31     : APPLICATION NUMBER: US 06/888,033
32     : FILING DATE: 18-JUL-1986
33     : PRIOR APPLICATION DATA:
34     : APPLICATION NUMBER: US 06/593,691
35     : FILING DATE: 26-MAR-1984
36     : ATTORNEY/AGENT INFORMATION:
37     : NAME: Smith, William M.
38     : REGISTRATION NUMBER: 30,223
39     : REFERENCE/DOCKET NUMBER: 12116-5-3
40     : TELECOMMUNICATION INFORMATION:
41     : TELEPHONE: 415-326-2400
42     : TELEFAX: 415-326-2422
43     : INFORMATION FOR SEQ ID NO: 2:
44     : SEQUENCE CHARACTERISTICS:
45     : LENGTH: 561 amino acids
46     : TYPE: amino acid
47     : TOPOLOGY: linear
48     : MOLECULE TYPE: protein
49     :
50     : US-08-358-901-2

```

Query Match	6.4%;	Score 127.5;	DB 1;	Length 561;
Best Local Similarity	23.2%;	Pred. No. 0.0004;		
Matches	83;	Conservative	44;	Mismatches 114;
				Indels 117;
				Gaps 21.

[illegible]

```

RESULT      3
US-08-566-347-2
: Sequence 2, Application US/08566347
: Patent No. 5633450
: GENERAL INFORMATION:
: APPLICANT: Suslow, Trevor V.
: APPLICANT: Jones, Jonathan D.G.
: TITLE OF INVENTION: No. 5633450el Chitinase-Producing Plants
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: One Market Plaza, Stewart Tower, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/566,347
: FILING DATE: 01-DEC-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/358,901
: FILING DATE: 19-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/550,253
: FILING DATE: 09-JUL-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 06/888,033
: FILING DATE: 18-JUL-1986
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 06/593,691
: FILING DATE: 26-MAR-1984
: ATTORNEY/AGENT INFORMATION:
: NAME: Smith, William M.
: REGISTRATION NUMBER: 30,223
: REFERENCE/DOCKET NUMBER: 12176-5-3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-326-2400
: TELEFAX: 415-326-2422
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 561 amino acids
: TYPE: amino acid
: TOPOLOGY: linear

```

MOLECULE TYPE: protein  
US-08-566-347-2

Query Match 6.4%; Score 127.5; DB 1; Length 561;  
Best Local Similarity 23.2%; Pred. No. 0.0004;  
Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

QY 2 NFTV---KSFVVICLLC---CLSTYVSIVY---EGHRRAPGESRKNP-REIITKTFK 48  
DB 99 NFKVKKGRYQMOV--ALCNADGCTASDATEIYVADTGSHP--LKEPLLEKKNPKYK 153  
QY 49 ESKGIIIGYPSWVSYNHNK-LDNPNLNVYHMSFAKMDLS---YDSIESIYGS-PL 101  
DB 154 QNSGKVVGSYFVEMGVGRNFTVDKIPQNLTHLYGFIPICGGNGINDSLKEIGSFQA 213  
QY 102 LKSLIGLEYIGLNEY-----FNDAMNLKRAPDIIMLSIG 138  
DB 214 LQSCQGRDEPKISIHDPFALQKQKGTAMDPPYKGNFGQLMALKQAHPLKILPSIG 273  
QY 139 GETYHPSSF-----DSALNAVERKIANLYDELGFDGIDVYE---PNSGFDGLNDREK 187  
DB 274 GWTLSDPFFMGDKVKRRFVGSVKEF--LQTKRFEDVDIDMERPGGKGNPNLGSFQD 331  
QY 188 ADFEYQVYTKLREYWCDDKLISISQSSNGALSCIGFNDPKKICMDEAPYNSKYFNKPDV 247  
DB 332 GETYVLLKKEILRAML--DQL-----SAETGRKY----- 357  
QY 248 KKEILRAAQMASGAIYLMNNLKMDIMVFYQTENYT-NSTDSTVMKELDSYAYYG 304  
DB 358 --ELTSA---ISAG-----KDKIDKV---AYNVAQNSMDHIFLM---SYDFYG 394

RESULT 4  
US-08-693-835-2  
Sequence 2, Application US/08693835  
Patent No. 5776448

GENERAL INFORMATION:  
APPLICANT: Suslow, Trevor V.  
APPLICANT: Jones, Jonathan D.G.  
TITLE OF INVENTION: No. 5776448el Chitinase-Producing Plants  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/693,835  
FILING DATE: 01-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/358,901  
FILING DATE: 19-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/550,253  
FILING DATE: 09-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/888,033  
FILING DATE: 18-JUL-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/593,691  
FILING DATE: 26-MAR-1984  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 12176-5-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 561 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-693-835-2

Query Match 6.4%; Score 127.5; DB 1; Length 561;  
Best Local Similarity 23.2%; Pred. No. 0.0004;  
Matches 83; Conservative 44; Mismatches 114; Indels 117; Gaps 21;

QY 2 NFTV---KSFVVICLLC---CLSTYVSIVY---EGHRRAPGESRKNP-REIITKTFK 48  
DB 99 NFKVKKGRYQMOV--ALCNADGCTASDATEIYVADTGSHP--LKEPLLEKKNPKYK 153  
QY 49 ESKGIIIGYPSWVSYNHNK-LDNPNLNVYHMSFAKMDLS---YDSIESIYGS-PL 101  
DB 154 QNSGKVVGSYFVEMGVGRNFTVDKIPQNLTHLYGFIPICGGNGINDSLKEIGSFQA 213  
QY 102 LKSLIGLEYIGLNEY-----FNDAMNLKRAPDIIMLSIG 138  
DB 214 LQSCQGRDEPKISIHDPFALQKQKGTAMDPPYKGNFGQLMALKQAHPLKILPSIG 273  
QY 139 GETYHPSSF-----DSALNAVERKIANLYDELGFDGIDVYE---PNSGFDGLNDREK 187  
DB 274 GWTLSDPFFMGDKVKRRFVGSVKEF--LQTKRFEDVDIDMERPGGKGNPNLGSFQD 331  
QY 188 ADFEYQVYTKLREYWCDDKLISISQSSNGALSCIGFNDPKKICMDEAPYNSKYFNKPDV 247  
DB 332 GETYVLLKKEILRAML--DQL-----SAETGRKY----- 357  
QY 248 KKEILRAAQMASGAIYLMNNLKMDIMVFYQTENYT-NSTDSTVMKELDSYAYYG 304  
DB 358 --ELTSA---ISAG-----KDKIDKV---AYNVAQNSMDHIFLM---SYDFYG 394

RESULT 5  
PCT-US94-01198-2

Sequence 2, Application PC/TUS9401198  
GENERAL INFORMATION:  
APPLICANT: Harman, Gary E.  
APPLICANT: Tronsmo, Arne  
APPLICANT: Hayes, Christopher K.  
TITLE OF INVENTION: Gene Encoding for Endochitinase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Specter, Eric S.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01198  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,269  
FILING DATE: 14-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Specter, Eric S.

REGISTRATION NUMBER: 22495  
REFERENCE/DOCKET NUMBER: CRF-D1060C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01198-2

Query Match 6.3%; Score 124.5; DB 5; Length 424;  
Best Local Similarity 21.0%; Pred. No. 0.00052;  
Matches 65; Conservative 51; Mismatches 93; Indels 101; Gaps 14;

QY 58 YPSPWSYNNHKLKLN-PNLNVHMSFAKMDL-SYDSTESIVGSPFLKSLGLEIYGLN 115  
DB 44 YFTNWGIYGRNFPQNLVASDITHVIYSFMNLEAYGVVVS-----GDAYADYQ 91  
QY 116 EYFND-----AMNLRKAPDIIMLSIGGETYHSPFSAL-----N 152  
DB 92 KHYDDSDWMDVGNNAVCYKQLEKLRKANGLKWLSTIGWTW-STNPPSASTDANKRN 150  
QY 153 AVEKRIANLVDELFGDIDVDEPEPNSFGDLNDEKADFFVQVYTKLRREYMCDDKLISIQ 212  
DB 151 FAKTATIFMKDWMGFDGIDVDW-----EYPADD-----TQ 179  
QY 213 SSNGALSCIGFNDPKKICMDEAPYNSKYFNKPDVKKELLRAQMASAGAIYLMNNLKD 272  
DB 180 ATNMVLL-----KEIRSQLDA-YAAQY-APGYHFLSLTAAPGPRHYSFLHMSDGLGQ 230  
QY 273 MIDMFVQTFNTNSTSTYMKELDYSAAYGKKYDYIINGFTLMFPSTFPNPNDK-M 330  
DB 231 VLDYVNLMAADYAGS-----WSSYS-----GHDAFLPANSNPNSSPYN 269  
QY 331 LKYSIGDFVK 340  
DB 270 TDQAIKDIYK 279

RESULT 6  
US-08-045-269C-2  
Sequence 2, Application US/08045269C  
Patent No. 5378821  
GENERAL INFORMATION:  
APPLICANT: Hartman, Gary E.  
APPLICANT: Tronsmo, Arne  
APPLICANT: Hayes, Christopher K.  
APPLICANT: Loritto, Matteo  
TITLE OF INVENTION: Gene Encoding for Endochitlinase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/045,269C  
FILING DATE: 14-APR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,784  
FILING DATE: 27-JUL-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/716,134  
FILING DATE: 17-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Spector, Eric S.  
REGISTRATION NUMBER: 22495  
REFERENCE/DOCKET NUMBER: CRF-D1060C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-045-269C-2

Query Match 6.1%; Score 120.5; DB 1; Length 424;  
Best Local Similarity 20.7%; Pred. No. 0.0013;  
Matches 64; Conservative 50; Mismatches 96; Indels 99; Gaps 14;

QY 58 YPSPWSYNNHKLKLN-PNLNVHMSFAKMDLSTDSIESIVGSPFLKSLGLEIYGLN 116  
DB 44 YFTNWGIYGRNFPQNLVASDITHVIYSFMNFOADG--TVVS-----GDAYADYQ 92  
QY 117 YFND-----AMNLRKAPDIIMLSIGGETYHSPFSAL-----NA 153  
DB 93 HYDDSDWMDVGNNAVCYKQLEKLRKANGLKWLSTIGWTW-STNPPSASTDANKRN 151  
QY 154 VEKRIANLVDELFGDIDVDEPEPNSFGDLNDEKADFFVQVYTKLRREYMCDDKLISIQS 213  
DB 152 AKTATIFMKDWMGFDGIDVDW-----EYPADD-----TQ 180  
QY 214 SSNGALSCIGFNDPKKICMDEAPYNSKYFNKPDVKKELLRAQMASAGAIYLMNNLKD 273  
DB 181 TNMVLL-----KEIRSQLDA-YAAQY-APGYHFLSLTAAPGPRHYSFLHMSDGLGQ 231  
QY 274 IDMFVQTFNTNSTSTYMKELDYSAAYGKKYDYIINGFTLMFPSTFPNPNDK-ML 331  
DB 232 LDYVNLMAADYAGS-----WSSYS-----GHDAFLPANSNPNSSPYNT 269  
QY 332 VKSIGDFVK 340  
DB 271 DQAIKDIYK 279

RESULT 7  
US-08-371-680-2  
Sequence 2, Application US/08371680  
Patent No. 6020540  
GENERAL INFORMATION:  
APPLICANT: Hartman, Gary E.  
APPLICANT: Tronsmo, Arne  
APPLICANT: Hayes, Christopher K.  
APPLICANT: Loritto, Matteo  
APPLICANT: Dipietro, Antonio  
TITLE OF INVENTION: Gene Encoding Endochitlinase  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/371,680  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,269  
FILING DATE: 14-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,784  
FILING DATE: 27-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/716,134  
FILING DATE: 17-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/184,115  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,390  
FILING DATE: 21-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,609  
FILING DATE: 15-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Spector, Eric S.  
REGISTRATION NUMBER: 22495  
REFERENCE/DOCKET NUMBER: CRF-D1424A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-371-680-2

Query Match 6.1%; Score 120.5; DB 3; Length 424;  
Best Local Similarity 20.7%; Pred. No. 0.0013;  
Matches 64; Conservative 50; Mismatches 96; Indels 99; Gaps 14;

QY 58 YPSVSVNHLKLDLN-PNLVNVHMSFAKMDLSYDSISYVSPILFSLGLEIYGLNE 116  
Db 44 YFTNMGATIGRNFOPNLVASDITHYISFPMNQADG--TVVS-----GDAYADYOK 92  
QY 117 YFND-----AMNLKARPDIIMLSLGETYHPSPFSAL-----NA 153  
Db 93 HYDDSDMVDGNNAVGCVKQKFLKLRANRLKMLSIGTWI-STNFSASASTDANRKNF 151  
QY 154 VEKIANLVDELGFPGIDVDYEPNGSFDGLNDEKADFFVQYVTKLREYMCDDKLISISQS 213  
Db 152 AKTATTEFKMDGFPDIDVDM-----EYPAAD-----TQA 180  
QY 214 SNGALSCIGFNDPKKICMDDEAPYNSKYENKPDYKELLRAQMASAGALYIMNLMNDM 273  
Db 181 TMMVLL-----KEINSQIDA-YAQT--APGYHFLSLIAPAPPEHYSELHMSDLCQV 231  
QY 274 IDWVFVQTFNTNSTDSTVMKELYSYAVYGRKKYDYVILMGFTLLFPSTPFPNDK--ML 331  
Db 232 LDYVNLMAVDYAGS-----WSSYS-----GHDNLNLFANPSNPSPINT 270  
QY 332 VKSIGDFVK 340  
Db 271 DQAIKDYIK 279

RESULT 8  
US-08-386-727-8  
Sequence 8, Application US/08386727  
Patent No. 5792647  
GENERAL INFORMATION:  
APPLICANT: ROSEMAN, SAUL  
APPLICANT: BASSLER, BONNIE

APPLICANT: KEYHANI, NEMAT O.  
APPLICANT: CHITLARD, EDITH  
APPLICANT: ROME, CHRIS  
APPLICANT: YU, CHARLES  
TITLE OF INVENTION: BACTERIAL CATABOLISM OF CHITIN  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/386,727  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOBBS, ANN S.  
REGISTRATION NUMBER: 36,830  
REFERENCE/DOCKET NUMBER: 4130/206916  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-386-727-8

Query Match 5.9%; Score 117.5; DB 1; Length 866;  
Best Local Similarity 21.9%; Pred. No. 0.0076;  
Matches 90; Conservative 58; Mismatches 140; Indels 123; Gaps 24;

QY 2 NETV---KYSEFLVILCLC---CLISTYVSI---EGHRRAPGESRRNP-REIITK 48  
Db 99 NEKVTKGRYQMQV--ALCNADGCTLSDKKEIYVADTDSHLAPLNA---PLQENKRPY 153  
QY 49 ESKGKITIGGYPSWSVSYNHLK-DLNPINLVVHMSFAKMDLS---YDSISYVS-PL 101  
Db 154 NKAGRVGAYVYEWGVYGRKFTVDKIPAKNLTHILYGFPTPGCGNGINDSLKEISGF 213  
QY 102 LFKSLGLEIYGLNEY-----FNDAMNLRKARPDIIMLSLG 138  
Db 214 LORSAGREDPEVSIHDPRAAVOMGQNLTADEPKKGFNGMLMALKRNPKILPSPG 273  
QY 139 GETY-HPSSFDALNAVEK-IANLVDELG---FDGIDVYE-----PNSGFDGLNDE 186  
Db 274 GWTLSDPFFYFEDTKRDTFVASKREYLOTWKFDPDVIDMEFPGOGANPILGPNDS- 332  
QY 187 KADFFVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMD-EPAPNSKYFKNP 245  
Db 333 ---ATYVALMKEL-----RAMLDELEAETGRQY--- 357  
QY 246 DYKRELLRAQMASAG--ATYLMNMLKMDIDWVFVQTFNTNSTDSTVMKELYSYAV 302  
Db 358 ---ELTSA---ISAGGDKIAKVYQAAQYNDYIFLMSYDSSGAPD---LKNLAHQTNL 407  
QY 303 YGKKYDYVILMGFTLFPSTPFPNDKMLVYSIGDFVYTKENTLNKRADGFG 353  
Db 408 YASSWD-----PATKYT-TDKGVKALLGGV--TPGKVVVVGAAAMYG 445

RESULT 9  
US-08-600-452A-8  
Sequence 8, Application US/08600452A  
Patent No. 5985644  
GENERAL INFORMATION:  
APPLICANT: ROSEMAN, SAUL  
APPLICANT: BASSLER, BONNIE  
APPLICANT: KEYHANT, NEMAT O.  
APPLICANT: CHITLARU, EDITH  
APPLICANT: ROWE, CHRIS  
APPLICANT: YU, CHARLES  
TITLE OF INVENTION: BACTERIAL CATABOLISM OF CHITIN  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & RICHARDSON P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/600,452A  
FILING DATE: 13-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07662/005001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 678-5070  
TELEFAX: (619) 678-5099  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-600-452A-8

Query Match 5.98; Score 117.5; DB 2; Length 866;  
Best Local Similarity 21.98; Pred. No. 0.0076;  
Matches 90; Conservative 58; Mismatches 140; Indels 123; Gaps 24;

QY 2 NFTV---KSFVLYICLLC---CLSTYVSVI---EGHRRPQESRKNP-REITKTFK 48  
DB 99 NFEVTVGGRQOMYV-ALCANADGTLSDKKEIVVADTDGSHLAPLNA---PLQENNKRYT 153  
QY 49 ESKGIIQGYPPSWVSNHNHL-DLNPNLNVHMSFAKMDL-----YDSIESIVGS-PL 101  
DB 154 NKAGKVVYGVYEWGVGKRFYVDKIPAKNLTHTLYGFTPGCGNGINDSLKEISGSPEA 213  
QY 102 LFKSLGLEITGLNEY-----FNDANLTKRARPDIIMLSLG 138  
DB 214 LQHSCHGREDFKVSINHPMAVOMGQGNLTADEPYKNGFNLMALKANPNLKLIPSVG 273  
QY 139 GETY-HPSSPDSALNAVEK-IANLVDELG---FDGIDVDYE-----PNSGFDGLNDE 186  
DB 274 GMLSDPFYFSDKTRKDTIVASMKELQTMKFPDGYDIDMEFPGGQGANNLGPRNDG- 332  
QY 187 KADFOVQYVTKREYMCDDKLISISQSSNGALSCIGFNDPKRICMD-EADYNSKYFNKP 245  
DB 333 ---ATYVALMKEI-----RAMUDELEAFGRQY----- 357  
QY 246 DYKELLRAQOMASAG---AIIYIMNMLKMDIMVFOVOTFYNTSTJSTVKELYDSIAY 302

DB 358 ---ELTSA---ISAGDKRIAKVDYQAQYMDYIFLMSYDFSGAFD---LKNLAHOTNL 407  
QY 303 YGKRYDYIIMGFILMPSTPFPNDKMLVKSIGDFVTEKLNKRADGFG 353  
DB 408 YASSWD-----PATKYT-TDKGVKALLGQV-TPEKVVYGAAMTG 445

RESULT 10  
US-08-392-625-24  
Sequence 24, Application US/08392625  
Patent No. 5837485  
GENERAL INFORMATION:  
APPLICANT: Enlian, Karl-Dieter  
APPLICANT: Gtz, Friedrich  
APPLICANT: Schnell, No. 5837485bert  
APPLICANT: Augustin, Johannes  
APPLICANT: Engelke, Gernar  
APPLICANT: Rosenstein, Ralf  
APPLICANT: Kaletta, Corina  
APPLICANT: Klein, Cora  
APPLICANT: Wieland, Bernd  
APPLICANT: Kupke, Thomas  
APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process For The Preparation  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,625  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652,0980002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2540  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 461 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-392-625-24

Query Match 5.68; Score 111; DB 2; Length 461;  
Best Local Similarity 21.08; Pred. No. 0.013;  
Matches 86; Conservative 50; Mismatches 125; Indels 148; Gaps 17;

QY 6 KYSFLVICLLCLL-----STYVSIVIGHRRPQESRKNPREITKTFKES 50  
DB 3 KEKFEIYFLISLVEFLONEXAFGSSLENEELSYSV-----EYDAAKTFKES 48  
QY 51 --GKGIQGY-----YPSWVSYHNLKDLNPNLN---VHMSFAKMDLsy 90  
DB 49 IKQNIETLTYPILHTAQTSTKSKLNSLIKSNKVKFVNPICSTCVKESVK----- 102

QY 91 DSIESIVGSPLEFKSLIGLEYIGLNEYFNDAMNLR-----KARP-----IMLSL 137  
Db 103 -----TGKNLNKK-----NGSHDLFDRQWDMRKITNEGSKYKLSPPKRAKVALYDS 150  
QY 138 GGETYHPSFSDSALNAVEKIANLVDELGPDGIDVD-----YE-----PN 176  
Db 151 GVNSSH-----IDLKSIINKIVEVRKNGRGSSENDSEGNKEEDKLNIGTLVAGOIGAN 205  
QY 177 GSFDDLNDKEKADFQVQYTKLREYWCDDK-LISISQSSNGALS-----CIGFNDPKKI 229  
Db 206 GNLKGVNPGVEANVYRVFSGSKSEMLWVSKGIIDAANDNDVYINSLGNYLLKDNQKKK 265  
QY 230 CMDDE-APYNSKYFNKPDYKKELLRAAQAASAGAIYIMNNLKMDIMVFOVTFNNTST 288  
Db 266 LRDEKDYDA-----LQKAINYAQKKGSIVVAAGNDGINVKKEINKKRNL 314  
QY 289 DSTVMKELDYSAIYKKYDYIIMGFTLMPSTPPNPKDKLYKSIGD 337  
Db 315 NSKTSKKYVS-----PANLNMMVTGVSIDD 340

RESULT 11  
US-08-466-961A-24  
; Sequence 24, Application US/08466961A  
; Patent No. 5843709  
; GENERAL INFORMATION:  
; APPLICANT: Entlan, Karl-Dieter  
; APPLICANT: G tz, Friedrich  
; APPLICANT: Schnell, No. 5843709bert  
; APPLICANT: Augustin, Johannes  
; APPLICANT: Engelke, German  
; APPLICANT: Rosenstein, Ralf  
; APPLICANT: Kaletta, Corina  
; APPLICANT: Klein, Cora  
; APPLICANT: Wieland, Bernd  
; APPLICANT: Kupke, Thomas  
; APPLICANT: Jung, G roland  
; APPLICANT: Kellner, Roland  
; TITLE OF INVENTION: Biosynthetic Process for the Preparation of  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,961A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/392,625  
; FILING DATE: 22-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/876,791  
; FILING DATE: 30-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/784,234  
; FILING DATE: 31-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 0652-0980004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 461 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-466-961A-24

Query Match 5.6%; Score 111; DB 2; Length 461;  
Best Local Similarity 21.0%; Pred. No. 0.013;  
Matches 86; Conservative 50; Mismatches 125; Indels 148; Gaps 17;

QY 6 KTSFLVLCICLL-----STYVSIEGHRARPGESKRNPREITKTFKS 50  
Db 3 KKKFIVFLILSLVLFQNEYARGSSLINEELSYYSV-----EYDNAKTFKS 48  
QY 51 --GKGIIQY-----YPSWVSYNHLLKDLNPLN--VYHMSPAKMDLSY 90  
Db 49 IKQNIETLYKIPELHTAOKTSKSLNLSKNKVKFVNPICSTCYVEKSYK----- 102  
QY 91 DSIESIVGSPLEFKSLIGLEYIGLNEYFNDAMNLR-----KARP-----IMLSL 137  
Db 103 -----TGKNLNKK-----NGSHDLFDRQWDMRKITNEGSKYKLSPPKRAKVALYDS 150  
QY 138 GGETYHPSFSDSALNAVEKIANLVDELGPDGIDVD-----YE-----PN 176  
Db 151 GVNSSH-----IDLKSIINKIVEVRKNGRGSSENDSEGNKEEDKLNIGTLVAGOIGAN 205  
QY 177 GSFDDLNDKEKADFQVQYTKLREYWCDDK-LISISQSSNGALS-----CIGFNDPKKI 229  
Db 206 GNLKGVNPGVEANVYRVFSGSKSEMLWVSKGIIDAANDNDVYINSLGNYLLKDNQKKK 265  
QY 230 CMDDE-APYNSKYFNKPDYKKELLRAAQAASAGAIYIMNNLKMDIMVFOVTFNNTST 288  
Db 266 LRDEKDYDA-----LQKAINYAQKKGSIVVAAGNDGINVKKEINKKRNL 314  
QY 289 DSTVMKELDYSAIYKKYDYIIMGFTLMPSTPPNPKDKLYKSIGD 337  
Db 315 NSKTSKKYVS-----PANLNMMVTGVSIDD 340

RESULT 12  
US-08-645-193B-26  
; Sequence 26, Application US/08645193B  
; Patent No. 5962253  
; GENERAL INFORMATION:  
; APPLICANT: Kupke, Thomas  
; APPLICANT: Gotz, Friedrich  
; APPLICANT: Kemper, Christoph  
; APPLICANT: Jung, Gunther  
; TITLE OF INVENTION: Oxidative Decarboxylation of Peptides  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/645,193B  
; FILING DATE: 13-MAY-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 0652.1540000

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 461 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 US-08-645-193B-26

Query Match 5.6%; Score 111; DB 2; Length 461;  
 Best Local Similarity 21.0%; Pred. No. 0.013;  
 Matches 86; Conservative 50; Mismatches 125; Indels 148; Gaps 17;

QY 6 KYSFLVLCCLL-----STYVYIEGHRARPGESRKNPREIITKTES 50  
 DB 3 KKEFFVFLILSLVFQNEAFAFGSSLNDELSTYSV-----EYDNKTFKES 48  
 QY 51 --GKGIIGY-----YPSWVSYNHNLKDLNPNLN--VYHMSFAKMDLSY 90  
 DB 49 IKRNIETLTIKIPELHTAQIKTSKSLNLSIKSNKVKFVNPCTCYEVSVK----- 102  
 QY 91 DSIESIVGSPLEKSLIGLEYIGLNEYFNDAMNR-----KARPD---IIMLSL 137  
 DB 103 -----TGKMLNNKK-----NGSHDLFDROWDMRKITNEGSKYKLSPRKAKVALVDS 150  
 QY 138 GGEYHPSSFDALNAVEKILANLYDELGFQIDVD-----YE-----PN 176  
 DB 151 GVNSH-----TDKSTINKIYNEVPKNGFSGENSEGNNKFEEDKLNHGLVLAQIGAN 205  
 QY 177 GSPDGLNDEKADFEVQYVTKLREYMCDDK-LISISOSNGALS-----CIGFNDPKI 229  
 DB 206 GNLKGVNPGVEAMNYRFGSKSEMLWVSGIIDAANDNDYIVNSGLNYLIKNNKK 265  
 QY 230 CMDEAPYNSKYFNKDYKKEILRAAOMASAGAIYILMNNKIDMVFQTFNYTNT 288  
 DB 266 LRDEKVDYDA-----LQKAINYAOKKGIYAAVANGNGINVKKKEIKKKNNL 314  
 QY 289 DSTVMEKLYSYAYGKKDYVILMFTLMPSTPFPNDKMLKYSID 337  
 DB 315 NSKTSKVIDS-----PAILNNMIVGSID 340

RESULT 13  
 US-09-025-691-3  
 Sequence 3, Application US/09025691  
 Patent No. 6069299  
 GENERAL INFORMATION:  
 APPLICANT: Broadway, Roxanne M.  
 APPLICANT: Harman, Gary E.  
 TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
 STREET: Clinton Square, P.O. Box 1051  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 14603  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/025,691  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/20120  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1600  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 376 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-025-691-3

Query Match 5.6%; Score 110.5; DB 3; Length 376;  
 Best Local Similarity 22.3%; Pred. No. 0.01;  
 Matches 68; Conservative 28; Mismatches 84; Indels 125; Gaps 15;

QY 33 PGESRKNPREIITKTESGKGIIGYYPWSVY--NHNLDL--NPNLNVYH----- 81  
 DB 2 PG---PGREKINL-----GYFTENGVRNTHVKNLVTSAGAEITHINISFGN 48  
 QY 82 -----SFAKMDLSYDSIESIVG-----SPLLFKSLIGLEYIGLNEYFNDAMNR 125  
 DB 49 VQGGKCTIGDSFAAYVDKAYTAESVDGADPTWDP-----LGRNPNQLKTK 95  
 QY 126 KARDIIMLSLGEYTHPSSFDALNA--VEKIANLYDELG---FDGIDVYE-PN- 176  
 DB 96 AKYPIKVLWSFGMTWSGGFTDAVKNPAFAKCHDLVEPRADYFDDIDIMEYRNA 155  
 QY 177 -----GSPDGLNDEKADFEVQYVTKLREYMCDDK-LISISOSNGALS-CIGFNDP 226  
 DB 156 GGLSDSGPAAKLNMQAMRAQGTDLVY-----AATFADASSGKL----- 198  
 QY 227 KVICMDEAPYNSKYFNKDYKKEILRAAOMASAGAIYILMNNKIDMVFQTFNYTNT 286  
 DB 199 -----DAADY-----AGAAQY-----FDWYNVMTYDFEG 222  
 QY 287 STDST 291  
 DB 223 AWDKT 227

RESULT 14  
 US-08-308-872B-6  
 Sequence 6, Application US/08308872B  
 Patent No. 5661006  
 GENERAL INFORMATION:  
 APPLICANT: BROWN, Thomas David Kay  
 APPLICANT: HORSBURGH, Brian Colin  
 TITLE OF INVENTION: CANINE CORONAVIRUS SUBUNIT VACCINE  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Akzo No. 5661006e1 Patent Department  
 STREET: 1300 Piccard Drive, Suite 206  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: U.S.A.  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/308,872B  
 FILING DATE: 19-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/042,846  
 FILING DATE: 05-APR-1993  
 PRIOR APPLICATION:



APPLICATION NUMBER: US 07/872,641  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 91.303.737.0  
FILING DATE: 25-APR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gormley, Mary E.  
REGISTRATION NUMBER: 34,409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1453 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Canine corona virus  
STRAIN: CCV-C54  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1453  
OTHER INFORMATION: /label= CCV-C54\_spike  
US-08-308-872B-6

Query Match 5.5%; Score 108.5; DB 1; Length 1453;  
Best Local Similarity 20.9%; Pred. No. 0.13; Indels 111; Gaps 17;  
Matches 78; Conservative 43; Mismatches 141;  
QY 10 LVICLLCCLSTVSVI-----EGHRAPEGSRKNPREIITKFEKSGKIIGYYP 60  
DB 1 MIVLCLCLLPSTSYNVCITSNNDQVAVVQLPGNENIKDFLEQNEEGSVVVGYP 60  
QY 61 SMVSN-----NHLKDLNPNLVNHSFAKMDISYSIESIVSPLLFK----- 104  
DB 61 TEVWNCSTRTATTAYHFSN-----IHAFYEDMEAMNSTGNARCKPLLVHVGSPV 113  
QY 105 SLIGLEYIGLMEYFNDAMNLRKRPDIIMLSIG-----GETVPSPSFDSA-----LN 152  
DB 114 SII-----VYSAIKDDVON---RP---LKHGLCLITKASTIDYNSFTSQWMDICLG 161  
QY 153 AVEKIANLVDELGEGIDVDEPNGS-EDGLNDEKADFFVOYVTKLEVMCDKLISIS 211  
DB 162 TDRKIP-----FSVYPTD---NGTKLEGL-----EMTDQYV---AYISDD-----S 197  
QY 212 QSSNGALSCIFNDPKKICMDDEAPYNSKYFNKPDYKKELLRAQMASAGAI----- 264  
DB 198 HRLN-----INTNMFNNVTIILYRSSTATWOKSAAYVYGVGSNF 236  
QY 265 --YLMNNLKDMDVYVOTFVNTSTSTVMKELYDVAAYGKKYDYIINGFTLMFPST 322  
DB 237 TYKLNNTNGLSKYEELCDIEYCTGATNVAPISGGIIPDGFSSNNFMILNLSSTFVSG 296  
QY 323 PFNPDKMLVRSI 335  
DB 297 RFTVQPLLVNCL 309

RESULT 15  
US-08-524-051-2  
Sequence 2, Application US/08524051  
Patent No. 5866788  
GENERAL INFORMATION:  
APPLICANT: Kramer, Karl J.  
APPLICANT: Muthukrishnan, Subbaratnam  
APPLICANT: Choi, Hee Kyung  
APPLICANT: Corpuz, Lolita  
APPLICANT: Gopalakrishnan, Bhuvana  
TITLE OF INVENTION: RECOMBINANT CHITINASE AND USE THEREOF AS  
TITLE OF INVENTION: A BIOCID

NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hovey, Williams, Timmons & Collins  
STREET: 2405 Grand Blvd., Suite 400  
CITY: Kansas City  
STATE: MO  
COUNTRY: USA  
ZIP: 64108  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/524,051  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Collins, John M.  
REGISTRATION NUMBER: 26,262  
REFERENCE/DOCKET NUMBER: 22875-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (816)474-9050  
TELEFAX: (816)474-9057  
TELEX: 434-363  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 554 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-524-051-2

Query Match 5.4%; Score 107; DB 2; Length 554;  
Best Local Similarity 20.4%; Pred. No. 0.041; Indels 28; Gaps 6;  
Matches 32; Conservative 40; Mismatches 57;  
QY 58 YPSPVWSY-----NHLKDLNPNLVNHSFAKMDLSYDSIESIVSPLIFKSLGLEXY 112  
DB 28 YFSNNAVYRPGVGRYGIIDI-PEVEKCTHIYISFICVTGNSVLLIIDPELDVCKNG--- 82  
QY 113 GLNEYFNDAMNLRKRPDIIMLSIGGETYHPSSF-----DSLNAVEKIANLVDELG 165  
DB 83 -----FRNFTSLRSHSPSVKFEVAVAGMAEGSSKYSHWAKSTMSFIRSVVSLKKYD 137  
QY 166 FDGIDVDYEPNGSFD---GLNDEKADFFVOYVTKLR 199  
DB 138 FDGLDLDMEYFGADRGGSFSDKR---FLIVYQELR 171

Search completed: May 8, 2002, 08:57:29  
Job time: 22 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 08:57:07 ; Search time 15.99 Seconds  
(without alignments)  
1800.749 Million cell updates/sec

Title: US-09-579-383-3

Perfect score: 1982

Sequence: 1 MNETVRYSEFLVTCILLCCLLS.....SDNAHNEQLATEYVESLH 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*

1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169.5	8.6	699	2	A38368
2	146.5	7.4	1088	2	D82246
3	140.5	7.1	610	2	JH0573
4	139	7.0	427	2	JC4565
5	136	6.9	274	2	JW0077
6	132.5	6.7	599	2	D83764
7	130.5	6.6	423	2	S51369
8	127.5	6.4	561	2	A25090
9	126.5	6.4	563	2	S60651
10	126.5	6.4	609	2	T42073
11	126.5	6.4	1054	2	T30933
12	124.5	6.3	511	2	S61166
13	123.5	6.2	417	2	T42074
14	122	6.2	291	2	S56698
15	121	6.1	285	2	S49898
16	120.5	6.1	424	2	S47133
17	119.5	6.0	831	2	T00323
18	119	6.0	466	2	T28216
19	119	6.0	558	2	T30418
20	117	5.9	276	2	S56696
21	116	5.9	291	2	S56697
22	115.5	5.8	504	2	A38221
23	114.5	5.8	413	2	JC2135
24	113	5.7	1213	2	T43916
25	112.5	5.7	571	1	JN0858
26	112.5	5.7	571	1	T42071
27	112.5	5.7	668	2	D70129
28	112.5	5.7	1051	2	D82428
29	112	5.7	492	2	F86868

30	111	5.6	285	2	S49879	hypothetical narbo
31	111	5.6	461	2	S23420	probable subtilisi
32	110.5	5.6	349	2	D70103	hypothetical prote
33	110.5	5.6	546	2	F84238	chitinase [importe
34	110	5.5	285	2	T12157	nodulin - fava bea
35	110	5.5	544	2	G64483	hypothetical prote
36	109.5	5.5	424	2	S68121	chitinase I precu
37	109.5	5.5	846	2	C82135	chitinase VC1952 I
38	107	5.4	554	2	A56596	chitinase (EC 3.2.
39	107	5.4	758	2	S37855	hypothetical prote
40	107	5.4	975	2	T08606	protein phosphatas
41	106.5	5.4	552	2	T41863	chitinase chl-A or
42	106	5.3	285	2	S49880	hypothetical narbo
43	105.5	5.3	736	2	B82944	ribose/galactose A
44	105	5.3	275	2	JE0184	chitinase (EC 3.2.
45	105	5.3	441	2	D83555	hypothetical prote

ALIGNMENTS

RESULT 1	
A38368	chitinase (EC 3.2.1.14) precursor - Bacillus circulans
C:Species: Bacillus circulans	
C:Date: 28-Jun-1991	#sequence_revision 28-Jun-1991 #text_change 15-Oct-1999
C:Accession: A38368	
R:Matnabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.	
J. Biol. Chem. 265, 15659-15665, 1990	
A:Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evol	
A:Reference number: A38368; MUID:90368776	
A:Accession: A38368	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-699 <MA>	
A:Cross-references: GB:M57601; GB:J05599; NID:q1066341; PIDN:AAA81528.1; PID:q142688	
C:Superfamily: fibronectin type III repeat homology	
C:Keywords: glycosidase; hydrolase; polysaccharide degradation	
Query Match	8.6%; Score 169.5; DB 2; Length 699;
Best local similarity	20.0%; Pred. No. 0.00028;
Matches	88; Conservative 58; Mismatches 113; Indels 181; Gaps 21;
QY	55 IGGYPSWVSY--NHNLDLPNVLVHMSFAKMDLSYDSIES----- 95
DB	46 IVGYPSMAVGRNVADIDPT-KVTHINAFADICWNGHGNPDSPGNPVWTQNE 104
QY	96 -----IVGSPLEFKSLIGLEYIG-----LNEYFDAMNLRKRPDIIMLSL 137
DB	105 KSQITVNGTIVLDPMID--TGKTFAGDTWDQPIAGINIQNLKQTNPNLTIISV 161
QY	138 GGETHPSFDSALNAV--EKIAN-LVDEL--GFGDIDVDE--PNSFGDLNREKAD 189
DB	162 GGMTSNRPSDVAATPAAREVFANSAVDPLRKRYNFDGVDLMEYVSGGLD-NSKRPED 220
QY	190 --FFQVYTKLEHYKCDPKLISISQSNGLSCIGFNDPKKICMDDEAPYNSKYRNKPDY 247
DB	221 KQNTVLLSKIREKL-----DAAGAV-----DG 243
QY	248 KKEILRAQMASAGATILMN---NLKMDIDMVFVQTFNTNSTDSIVMKEL---YDSY 300
DB	244 KKYLL--TIASGASATYAANTELAKIAIYDWMIMMYDFGAWQKISAHNAPLNYDPA 300
QY	301 A-----YGGKKYD----- 308
DB	301 ASAGVDPANTFNVAAGQCHLDAGVPAKLVLYGPFYGRGMDGCAQAGNGOYOTCTGGS 360
QY	309 -----YVIMGFTLMFPSTP-----FNPDKMLV-----KSIQDEYK 340
DB	361 SVGTWEAGSPFDYDEANVYINKNGYTRWMDPAKVPYILNMSNKRFFISYDAESVG--YK 418
QY	341 TENKLNKRAADGGLMSLSSD 360

Db 419 TAYIKSKLGGAMFWELSGD 438

## RESULT 2

probable chitinase VC1073 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: D82246

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bess, S.; Qin, H.; Dragol, I.; Sellers, F.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: D82246

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1088 <HE1>

A:Cross-references: GB:AE004188; GB:AE003852; NID:9655530; PIDN:AAF94232.1; GSPDB:GN001

C:Genetics: Experimental source: serogroup O1; strain N16961; biotype El Tor

A:Gene: VC1073

A:Map position: 1

Query Match 7.4% Score 146.5; DB 2; Length 1088;

Best Local Similarity 23.6%; Pred. No. 0.021; Mismatches 103; Indels 93; Gaps 21;

Matches 79; Conservative 60; Mismatches 103; Indels 93; Gaps 21;

Qy 39 NPREIITFKESGKGIIOGYPSWVSYNHNK-----DLNPLNVVHNSFA-----KMDL 88

Db 383 HPRRLI-----GYTSMRTGKNGLPAYLAGDL-PWEKTLTHINYPASINKSDF 429

Qy 89 SYDSIEIVSPLFKSLIGLE-----YIGLNEYFNDAMLRKRPDIIMLSIGG--E 140

Db 430 SMOVDSD--ATKMTWENVPGAEKMPSLPYQG--HFNLSEFKQYPPVKTLISVGWAE 484

Qy 141 T--YHPSSPSFA-----LNAYEKIA-NLYDELGPDGIVDYE-PGSPFDGLDKKA 188

Db 485 TGGFYPTTTDLASCYVNMKGKAFNKSADVDFIRQDFGVDIDYEPSSMKDSGN---PV 541

Qy 189 DFEVQVYTKLREYMCDDKLISISOSNGALSCIGFNDPKKICMDEAFYNSKYFNKPPVK 248

Db 542 DF--EQSNGKRCGQLMDNVMVMTLRK-ALDKAGEEDGRRLMTLASP--SSAY----- 590

Qy 249 KELLRAAQMASAGAITLNNLKMIDMVFYQTFN-----YTNSIDSTVM 293

Db 591 --ELRGKQDFA-----MDVLDYVNIIMSYDLHGTWNEFVGPOALFPDDGDADLA 638

Qy 294 K-ELYDSYAYGKKYDVYIIMGFTLMPSTPFPN 327

Db 639 KMGVYTTAEYQIGY---LNOAMTHHFFRCAGFKPS 670

RESULT 3

chitinase (EC 3.2.1.14) 63 - Streptomyces plicatus

C:Species: Streptomyces plicatus

C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 31-Mar-2000

C:Accession: JH0573; A29912

R:Robbins, P.W.; Overbye, K.; Albright, C.; Benfield, B.; Pero, J.

Gene 111, 69-76, 1992

A:Title: Cloning and high-level expression of chitinase-encoding gene of Streptomyces p

A:Reference number: JH0573; MUID:92192480

A:Accession: JH0573

A:Molecule type: DNA

A:Residues: 1-610 <ROB>

A:Cross-references: GB:M82804; NID:q153215; PIDN:AAA26720.1; PID:q153216

A:Note: the authors translated the codon AAG for residue 360, 406, 431, 442, 460, 489, 4

R:Robbins, P.W.; Albright, C.; Benfield, B.

J. Biol. Chem. 263, 443-447, 1988

A:Title: Cloning and expression of a Streptomyces plicatus chitinase (chitinase-63) 1

A:Reference number: A29912; MUID:88087127

A:Accession: A29912

A:Molecule type: DNA

A:Residues: 1-2, '1', 4-45 <RO2>

A:Cross-references: GB:M18397; NID:q153208; PIDN:AAA26717.1; PID:q153209

C:Comment: This protein is related with the steps in chitin utilization.

C:Genetics:

A:Gene: chitA

C:Superfamily: Streptomyces plicatus chitinase 63; bacterial cellulose-binding domain

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:33-132/Domain: Bacterial cellulose-binding domain homology <BCB>

Query Match 7.1% Score 140.5; DB 2; Length 610;

Best Local Similarity 21.4%; Pred. No. 0.026;

Matches 79; Conservative 51; Mismatches 108; Indels 131; Gaps 17;

Qy 34 GESRRNPREIITFKESGKGIIOGYPSWVSY--NHNKDL-----NPNLN--- 77

Db 232 GEPNPNP-----GAEYKMGFTWVGYYGRNHYHVKNLVTSASAKITHINLRFQNV 281

Qy 78 -----VYHMSFAKMDLSDSIESTVG-----SPLFKSLIGLEYIGLNEYFNDAMLRK 126

Db 282 OGGKCTIDADAYDYKAYTADQSVGVADPTWDP-----LRANFQRLNKA 328

Qy 127 ARPDIMLSIGGFTY-----HPSSFDALNAVEKIANLYDELGPDGIVDYE-PN 176

Db 329 EYPIHKLILYSGGWTWSGFPDPAKPNPAFAKSGHDLVEDPRMADV--FGIDIDMWEYRN 366

Qy 177 GSFDDLNDKERADFEVQVYTKLREYMCDDKLISISOSNGALSCIGFNDPKKICMDEAP 236

Db 387 AGLSCDETSAPNAFSSMKAMRAEFQGDYLTITAAVADG-----SDGKI---DAAD 436

Qy 237 YN--SKYFNKPDYKELLRAAQMASAGAITLNNLKMIDMVFYQTFNNTNSDSTVM 294

Db 437 YGEASKY-----IDVNMWYTFD----- 454

Qy 295 ELYDSYAYYG-----KKYDVYIIMGFTLMPSTPFPN-----PNDKMLVKSIGDFYKTE 342

Db 455 --FGAMAKNGPAPHPSPINATVDGIPQSGFTTADMAAFKSKGVADKLLI-GIGFYGRGW 511

Qy 343 NKLNRADG 351

Db 512 TGVTSAPG 520

RESULT 4

JC4565

chitinase (EC 3.2.1.14) 1 precursor - Coccidioides immitis

N:Alternate names: complement fixation antigen homolog; CTS1 protein

C:Species: Coccidioides immitis

C>Date: 12-Mar-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Nov-1998

C:Accession: JC4565

R:Pishko, E.J.; Kirkland, T.N.; Cole, G.T.

Gene 167, 173-177, 1995

A:Title: Isolation and characterization of two chitinase-encoding genes (cts1, cts2)

A:Reference number: JC4565; MUID:96144270

A:Accession: JC4565

A:Molecule type: mRNA

A:Residues: 1-427 <PTS>

A:Cross-references: GB:L41663

A:Experimental source: C735

C:Genetics:

A:Gene: cts1

A:Introns: 47/3; 171/3; 191/3; 215/2; 391/3

C:Superfamily: Streptomyces chitinase ch140

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-17/Domain: signal sequence #status predicted <Sig>

F:18-427/Product: chitinase 1 #status predicted <Mat>

F:387/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.0%; Score 139; DB 2; Length 427;  
Best Local Similarity 23.7%; Pred. No. 0.021;  
Matches 89; Conservative 55; Mismatches 118; Indels 114; Gaps 21;

OY 9 FLVICLCCLSTYYSVIEGHRARGESRKNREIIKFEKSG-KGIIGGYPSWVS-- 65  
DB 3 FLIGALL-TLQTLV-----OASSMSMPNSYPVEADPAEGGFSVY--YVWNAIYGR 52  
OY 66 NNLKDLNPLNVVHMSFA-----KMDLSYDSIESIVGSPFLKSLGLEIYL 114  
DB 53 GNNPDQLKAD-QFHILYLFANIRSGEYLLSDTMADTKKHFGDK-----MDERPN 103  
OY 115 NEY--FNDAMLRKARPDIIIMLSIGETYPHSSDSALNAVE---KIAN---LVDELG 165  
DB 104 NYGCGIKQWYLLKKNRNKTLISIGCWYSP-NFKTPASTEGRKKFADTSLKIMKDG 162  
OY 166 FPGIDVDE-PNGSPDGLDKKADFEYVYTKIREYMDOKLISISOSNGALSCIGN 224  
DB 163 FPGIDIDWYTP-----EDERQANDFVLL----- 186  
OY 225 DPKKICMDDEAFYNSKYENKPDVKELRAAOMASAGAIYLMNMLKDM---IDMWYVQT 281  
DB 187 ---KACREALDAYSAKH---FNGKKFLTLTA--SPAGQNYKKLKLAEKDYLDLFWNLMA 238  
OY 282 FNYTSTD--STVMKELDYSAIYGGKYDYIIMGFTLMPSTPFPNDKMLVKSIGDFV 339  
DB 239 YDFSGSMDKSVSGHMSNVPS-----TKPESTPFSSD-----KAVKDYI 277  
OY 340 KTEENKLNKRAODGFLM 355  
DB 278 KAGVPANKIYLGMPLY 293

RESULT 5  
JM0077  
chitinase (EC 3.2.1.14) a - Gladiolus gandavensis  
C:Species: Gladiolus gandavensis  
C:Date: 17-Jun-1998 #sequence\_revision 10-Jul-1998 #text\_change 07-May-1999  
C:Accession: JM0077  
R:Yamagami, T.; Mine, Y.; Ishiguro, M.  
Biosci. Biotechnol. Biochem. 62, 386-389, 1998  
A:Title: Complete amino acid sequence of chitinase-a from bulbs of gladiolus (Gladiolus  
A:Reference number: JM0077; MUID:98193995  
A:Accession: JM0077  
A:Molecule type: protein  
A:Residues: 1-274 <YAM>  
A:Experimental source: bulbs  
C:Comment: This enzyme hydrolyzes beta-1,4-linked N-acetylglucosamine polymer chitin.  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: glycosidase; hydrolase

Query Match 6.9%; Score 136; DB 2; Length 274;  
Best Local Similarity 22.4%; Pred. No. 0.019;  
Matches 74; Conservative 49; Mismatches 102; Indels 106; Gaps 16;

OY 72 LNPINLVH--MSFAKMDLSYDSIESIVGSPFLKSLGLEIYGNEXFN--DAMNLRK 126  
DB 21 INPRITDFQVLSFA-----IDYVSTP-----HTPFGKFNWVMDSSNLGP 62  
OY 127 ARPDII-----MLSLGET-----YHSSFDSNL-NAVEKIANLVDELGFDGID 170  
DB 63 SOVDIKSHNNVRAVSLGATVGKSVQPSIDSMVWNAVSLQIIDRYMLDGD 122  
OY 171 VDYPNGSFDGLNDEKEDFEYVYTKIREYMDOKLISISOSNGALSCIGFNDPKIC 230  
DB 123 IDYENFQNTD-----PDTRAECTIGRLITLAKRNVINFAI----- 158  
OY 231 MDDEAPYNSKYFNKPDVKELRAAOMASAGAIYLMNMLKDMIDMVFVQFNNTSIDS 290  
DB 159 ---APF-----PDVEEYLA-----LMNRKYNVINHINQFAYADSSSTTV 195  
OY 291 TVMKELIYSYA--YGGKYDYIIMGFTLMPSTPFPNDKMLVKSIGD---FVTEENKL 345

DB 196 RQFLKXYDPAVSKYRG-----GNVLISFSTERDAGLTFVDRGFPAASILKRQKRL 246  
OY 346 NKRAOGFLMS--LSSDNAHNEOLAIEYFV 374  
DB 247 H-----GIAVMSADTSKSNGRFIDEAOSFLV 273

RESULT 6  
DB3764  
chitinase BH0916 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: DB3764  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: AB3650; MUID:20263314  
A:Accession: DB3764  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-599 <STO>  
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BAB04635.1; GSPDB:G  
C:Genetics:  
A:Gene: BH0916

Query Match 6.7%; Score 132.5; DB 2; Length 599;  
Best Local Similarity 22.2%; Pred. No. 0.094;  
Matches 62; Conservative 42; Mismatches 82; Indels 93; Gaps 13;

OY 55 IOGYPPSWVS--NNMLKDLNPLNVHMSFAKMDLSYDSIES----- 95  
DB 42 IVAYYPSMGAYGRDYQVMDIDAS-KISHINYAFANICWDRGHPDPAGPNQTWSCODE 100  
OY 96 -----IVGSPLL--FKSLGLEI-IGLNEFNDAMLRKARPDIIIMLSIGET 141  
DB 101 NGVIDVPNGSIYMGDPWIDAKRSNPGDTWDEPLRGNFKOLNKTKEHPHLKTLISYGVWT 160  
OY 142 YHSSFDSNL--NAVEKIANLVDELGFDGIDVDE--PNGSPDGLN---DKKAD 189  
DB 161 WSNRSDMAATETRNFPANSVAEFTKRYGFDGVYDWMYPSGGLPGNSRREDKENNY 220  
OY 190 FEYVYTKIREYMDOKLISISOSNGALSCIGFNDPKKICMDDEAPYNSKYFNKPDVK 249  
DB 221 LLLQEV-----DKL-----DEA-----GQEDGRD 240  
OY 250 ELLRAAOMASAGAIYLMN-----LKDMIDMVFVQFN 284  
DB 241 YLLTIASGASPG---YVENNKLNEIAEIVDWINIMTYDF 276

RESULT 7  
S51369  
chitinase - fungus (Trichoderma harzianum)  
C:Species: Trichoderma harzianum  
C:Date: 15-Jul-1995 #sequence\_revision 19-Apr-1996 #text\_change 28-May-1999  
C:Accession: S51369  
R:Garcia, I.; Lora, J.M.; de la Cruz, J.; Benitez, T.; Lobell, A.; Plincor-Toro, J.A.  
Curr. Genet. 27, 83-89, 1994  
A:Title: Cloning and characterization of a chitinase (CHT42) cDNA from the mycoparas  
A:Reference number: S51369; MUID:95269313  
A:Accession: S51369  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-423 <GAR>  
A:Cross-references: GB:S78423; NID:g999375; PIDN:AAB34355.1; PID:g999376  
C:Superfamily: Streptomyces chitinase chl40

Query Match 6.6%; Score 130.5; DB 2; Length 423;  
Best Local Similarity 21.6%; Pred. No. 0.083;





C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Jun-2000  
C:Accession: T42074  
R:Saito, A.; Fujii, T.; Yoneyama, T.; Redenbach, M.; Ohno, T.; Watanabe, T.; Miyashita, submitted to the EMBL Data Library, August 1998  
A:Description: Structure of seven different chitinase genes from Streptomyces coelicolor  
A:Reference number: Z22056  
A:Accession: T42074  
A:Status: preliminary; translated from GB/EMBL/DDaj  
A:Molecule type: DNA  
A:Residues: 1-417 <SAI>  
A:Cross-references: EMBL:AB017011; PDB:BA075645.1  
C:Gene: chid  
C:Superfamily: Streptomyces chitinase chl40

Query Match 6.2%; Score 123.5; DB 2; Length 417;  
Best Local Similarity 23.0%; Pred. No. 0.25;  
Matches 64; Conservative 42; Mismatches 71; Indels 101; Gaps 16;

QY 55 IGGYPSWVSY--NNHKLKDLNPN-----LNVMHMSF-----AKMDSYDSI 93  
DB 55 VGGYFEMGVYDKNHYNHKNIESSGADKLTHINISFGNVTGGKCAMGDVATRAITAA 114  
QY 94 ESTIVG-----SPLFKSLIGLEYIGLNEYFNAMNLRKARPDIIIMLSIGETV---- 142  
DB 115 DSDVGVADPTWQDP-----LRGNPNOLKRLKQHPDLKILMSFGMTWSGGFA 161  
QY 143 ----HSSSF-DSALNVE--KIANLVDLGFDDGIDVYF--PNSGFDLNL-DKEADFFVQ 193  
DB 162 QAAQNEPAPQSCYDILVENSRRMDV----FDGIDIMWEYFNAC--GISCTSGRDAFPK 214  
QY 194 VYTKLEMYMCDKLISISOSSNGALSCIGFNDPKICMDDEAPYNSKYFNPDVKKELLR 253  
DB 215 LMGALNAKGGQDYLVAIRAADATAG-----GKI---DAADV----- 248  
QY 254 AAOMASAGGATYLMNNLKMIDMVFOTFNTYSTST 291  
DB 249 -----AGAAQY-----VDWYNPMTYDFPGAMDAT 272

RESULT 14  
S56698  
narbonin (clone pNag2) - *Vicia narbonensis*  
N:Alternate names: 2S globulin  
C:Species: *Vicia narbonensis*  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Nov-1999  
A:Accession: S56698; S56735; S44031  
R:Nong, V.H.; Schlegel, B.; Bassuener, R.; Repik, A.; Horstmann, C.; Muentz, K.  
Plant Mol. Biol. 28, 61-72, 1995  
A:Title: Narbonin, a novel 2S protein from *Vicia narbonensis* L. seeds: cDNA, gene structure  
A:Reference number: S56696; MUID:95306792  
A:Accession: S56698  
A:Molecule type: DNA  
A:Residues: 1-291 <NON>  
A:Cross-references: EMBL:Z25533; NID:9396824; PDB:CAA80980.1; PID:9396825  
A:Accession: S56735  
A:Molecule type: protein  
A:Residues: 2-15:182-191, 'E', 193-195, 'EE', 198-207, 250-251, 'E', 253-261, 290-291 <NOW>  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: seed; storage protein  
E:2-291/Product: narbonin #status experimental <MAT>

Query Match 6.2%; Score 122; DB 2; Length 291;  
Best Local Similarity 21.5%; Pred. No. 0.2;  
Matches 72; Conservative 39; Mismatches 84; Indels 140; Gaps 18;

QY 72 LNPNLNVVMSFAKMDLSYDSIVGSPLL-FKSLIGLEYIGLNEYFN----- 120  
DB 12 VKPNSTTLH-----DSPREIVNTETLEPHYILG---FAIESIYESGKGTGTFEES 58

QY 121 -----AMNLKRAPDIIMLSIGETVHPSFSDA-----LNAVEKIANTLV--- 161  
DB 59 MDVELFGEKVKNLKRHHPEKVVVISIGGVN--TPEDPAEENWVMSNAKESKLIIQYK 117  
QY 162 -DEIG--FDGIDVYFENGSPDGLNDEKADFFVQYTKLRIMCDKL-----ISISQSS 214  
DB 118 SDDSGNLIDGIDIHIEHRS-----DEFPATLMGQLLTLELK---DDDLINIVVSIAPSE 169  
QY 215 NALSCIGFNDPKICMDDEAPYNSKYFNKPKVKKELLRAAOMASAGATYLMNNLKMID 274  
DB 170 N-----NSSHYK-----LYNAKKDYI 186  
QY 275 DMVFVGFNTN--STDST--YMKELYSYAYYKRYDVIITG----- 315  
DB 187 NMVDYGFSSNOOKPVSTDAEVEIFKSLEKDY-----HPHKVLPGFSTPLDKHNRK 240  
QY 316 -----TLMPSTPF-NPNDKMLYKISGD 337  
DB 241 DIFIGCTRLVQTFSLPGVFEMANDSVIPKRDGD 275

RESULT 15  
S49898  
hypothetical narbonin-like 2S protein (clone pYSNAG2) - spring vetch  
C:Species: *Vicia sativa* (spring vetch, tare)  
C:Date: 05-Mar-1995 #sequence\_revision 14-Jul-1995 #text\_change 23-Mar-2001  
A:Accession: S49898  
R:Nong, V.; Muentz, K.  
submitted to the EMBL Data Library, November 1994  
A:Description: A genomic sequence encoding putative narbonin from *Vicia sativa*.  
A:Reference number: S49880  
A:Accession: S49898  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <NON>  
A:Cross-references: EMBL:Z46835; NID:9600112; PID:9600113  
C:Superfamily: alcohol sulfotransferase

Query Match 6.1%; Score 121; DB 2; Length 285;  
Best Local Similarity 21.1%; Pred. No. 0.23;  
Matches 60; Conservative 55; Mismatches 89; Indels 80; Gaps 16;

QY 114 LNEYF--NDAMNLRKAPDIIMLSIGETVHPSFSDA-----LNAVEKIANTLV--- 161  
DB 59 MDEFPGDPKYNLKTKEPEKVVVISIGRGVE--TPDPAPQNIWVSNVSKLIIQYK 117  
QY 162 DEIG--FDGIDVYFENGSPDGLNDEKADFFVQYTKL-REYMCDDKLISISQSSNGAL 218  
DB 118 NESGNLIDGIDINIEHRS-----DEAFPRILGQLITELKENDLNHVVSIAPSENNAS 172  
QY 219 SCIGFNDPKICMDDEAPYNSKYFNKPKVKKELLRAAOMASAGATYLMNNLKMIDMVF 278  
DB 173 SYLNL-----YNA-----NPDDI-----NLVDYGF 192  
QY 279 VQTFNNTNSDGS--YMKELYSYAYYKRYDVIITGFTLMPSPDPNDKMLYKISG 336  
DB 193 SNOLRHVSTEDAPDYIKRVVNDY-----FTHKVLPGFST-----DPLNMMTKITR 239  
QY 337 D-FYKTKNKKLRAD--GFGWLSLSDGNMAHQALIEFVSE 377  
DB 240 DIFIGCTRLKQTSLSLPGVFEMANDSVIPKRDGD 275

Search completed: May 8, 2002, 08:58:23  
Job time: 76 sec



---



GenCore version 4.5  
Copyright (c) 1993 - 2000. Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 8, 2002, 08:58:02 ; Search time 13.25 Seconds  
(without alignments)  
1045.985 Million cell updates/sec

Title: US-09-579-383-3  
Perfect score: 1982  
Sequence: 1 MNPTVKYSPFLVTCILCLLS.....SDNAHNEQLAIEFYVESLH 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%.

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169.5	8.6	699	1 CH11_BACCI	P20533 bacillus ci
2	140.5	7.1	610	1 CH1T-STRL	P11220 streptomyc
3	132.5	6.7	427	1 CH1L-COCIM	P54196 coccidioid
4	130.5	6.6	423	1 CH1A-TRHA	P48827 trichoderm
5	128.5	6.5	563	1 CH1A-SERVA	P07254 serratia ma
6	126.5	6.4	619	1 CH1T-STRLI	P36909 streptomyc
7	115.5	5.8	504	1 CH1T-BRUMA	P28030 brugia mala
8	111	5.6	461	1 BRIP-STARP	P30199 staphylococ
9	110.5	5.6	349	1 Y028-BORBU	O51059 borrelia bu
10	110	5.5	544	1 YE72-METUA	O58867 methanococ
11	108	5.4	587	1 DRTS-PLABA	O27713 plasmodium
12	107	5.4	758	1 CH1T-MANSE	P36362 manduca sex
13	107	5.4	758	1 YKD4-YEAST	P36096 saccharomyc
14	105	5.3	492	1 PD1L-SCRHO	O10057 schizosacch
15	104.5	5.3	617	1 CH1T-CAREL	O11174 ctenothadri
16	104	5.2	1401	1 LAT4-LATWA	P23631 latrodectus
17	102.5	5.2	430	1 AST2-YEAST	P39945 saccharomyc
18	102.5	5.2	493	1 Y130-MYCPN	P75506 mycoplasma
19	102	5.1	327	1 VP26-MOUSE	P40336 mus musculu
20	101.5	5.1	1046	1 CH1D-VIBU	P66156 vibrio furn
21	100.5	5.1	423	1 CH1L-APHAL	P32470 aphanoeladi
22	100.5	5.1	597	1 CH1X-STROI	O05638 streptomyc
23	100.5	5.1	785	1 VD05-VACCC	P21010 vaccinia vi
24	100.5	5.1	822	1 DEXT-STRLS	O59979 streptococc
25	100	5.0	1139	1 LY14-YEAST	O59024 methanococc
26	99.5	5.0	790	1 PR22-YEAST	P40971 saccharomyc
27	99.5	5.0	1145	1 PR22-YEAST	P24386 saccharomyc
28	99.5	5.0	1323	1 ADRI-YEAST	P07248 saccharomyc
29	99	5.0	314	1 LEP-BUCAL	P57347 buchnera ap
30	98.5	5.0	445	1 GLIM-HELPI	P51747 helicobacte
31	98.5	5.0	537	1 SYR-MYCE	P47618 mycoplasma
32	98.5	5.0	1448	1 UGGG-SCRHO	O09140 schizosacch
33	98	4.9	327	1 VP26-HUMAN	O75436 homo sapien

34	97.5	4.9	539	1 OGP-SHEEP	Q28542 ovvis aries
35	97	4.9	550	1 CH1T-NPVOP	O10363 orgyia pseu
36	97	4.9	551	1 CH1T-NPVOP	P41684 autographa
37	97	4.9	986	1 EP1B-STARP	P30195 staphylococ
38	97	4.9	1186	1 CAGA-HELPI	P55980 helicobacte
39	96.5	4.9	483	1 Y045-MYCE	P47291 mycoplasma
40	96.5	4.9	785	1 VD05-VACCV	P04305 vaccinia vi
41	96.5	4.9	2710	1 TOXA-CLADI	P16154 clostridium
42	96	4.8	454	1 AATM-LUPAN	P26563 lupinus ang
43	96	4.8	1018	1 YC14-METUA	O58611 methanococ
44	96	4.8	1082	1 RRPO-ROTPC	P26190 porcine rot
45	95.5	4.8	671	1 OGP-MESAU	Q60557 mesocricetu

## ALIGNMENTS

```

RESULT 1
CH11_BACCI STANDARD; PRT; 699 AA.
AC P20533;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE CHITINASE A1 PRECURSOR (EC 3.2.1.14).
GN CH1A1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1387;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WL-12;
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
RT its evolutionary relationship to Serratia chitinase and to the type
RT III homology units of fibronectin."
RL J. Biol. Chem. 265:15659-15665(1990).
RN [2]
RP MUTAGENESIS.
RC STRAIN-WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
RT chitinase activity."
RL J. Biol. Chem. 268:18567-18572(1993).
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M57601; AAA81528.1; -.
CC PIR: A38368; A38368.
CC HSSP: P02751; 1FTG.
CC InterPro: IPR003610; Chitin_bind3.
CC InterPro: IPR001579; Chitinase_2.
CC InterPro: IPR001777; FN_III.
CC InterPro: IPR001223; Glyco_hydro_18.
CC Pfam: PF000041; fn3_2.
CC Pfam: PF00704; Glyco_hydro_18; 1.
CC PRINTS: PR00014; FNTPETIT.

```



Db 437 YGASKY-----IMWYNTYDF----- 454

OY 295 ELYSDYAYG-----KKDYVIMGFTLMFPSTPPN---PNDKMLVSGDFVTE 342

Db 455 --FGMAWNGPAPSPPLNADYDGIQOGFTTRADAMAKRSKGVPADKILLI-GIGYGGW 511

OY 343 NKLNRKADG 351

Db 512 TGVTSAPG 520

RESULT 3

CHIL\_COCIM STANDARD; PRT; 427 AA.

ID CHIL\_COCIM

AC P54196;

DT 01-FEB-1996 (Rel. 34, Created)

DT 01-FEB-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE ENDOCHITINASE 1 PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN) (CF-ANTIGEN) (CF-Ag).

GN CTS1.

OS Coccidioides immitis.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Onygenales; mitosporic Onygenales; Coccidioides.

OX NCBI\_TaxID=5501;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C735;

RX MEDLINE=96144270; PubMed=8566773;

RA Pishko E.J., Kirkland T.N., Cole G.T.;

RT "Isolation and characterization of two chitinase-encoding genes (cts1, cts2) from the fungus Coccidioides immitis.";

RL Gene 167:173-177(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=SLIVEIRA;

RA Yang C., Zhu Y., Magee D.M., Cox R.A.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.

CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: LA1663; AAA92643.1; -

DR EMBL: U51271; AAA96515.1; -

DR EMBL: U33265; AAB06687.1; -

DR HSSP: P07254; ICTN.

DR InterPro: IPR001579; Chitinase\_2.

DR InterPro: IPR001223; Glyco\_hydro\_18.

DR Pfam: PF00704; Glyco\_hydro\_18; 1.

DR PROSITE: PS01095; CHITINASE\_18; 1.

KW Hydrolyase; Glycosidase; Chitin degradation; Chitin-binding; Signal;

KW Glycoprotein.

FT SIGNAL 1 ?

FT CHAIN ?

FT CARBOHYD 387 427

FT CONFLICT 15 47

FT CONFLICT 199 199

FT SEQUENCE 427 AA; 4/629 MW; IC396DBDBIA/001A CRC64;

Query Match 6.7%; Score 132.5; DB 1; Length 427;

Best Local Similarity 25.3%; Pred. No. 0.039;

Matches 62; Conservative 39; Mismatches 65; Indels 79; Gaps 12;

OY 124 LRKAPDIIIMLSIGETYPHSPSSPDAIYNNKDM---IDWVFQTFNTNSTD--S 290

Db 115 LKNNRNKLTLLSIGWYSP-NFKTPASTEGRKKFPDTSKLKKDLGFDGIDWMEY 173

OY 176 NGSFDGLNDEKADFFVOYVTKLRKREYCDKLTISQSSNGALSCIGFNDPKKICMDDEA 235

Db 174 -----EDEKQANDFVLL-----KACREALD 194

OY 236 PYSKATFNKPDYKKELELRAQWASAGAIYLMNNLKDM---IDWVFQTFNTNSTD--S 290

Db 195 AYSAKH---PNSKFLTLTA--SPAGPQYNNKLAEMDKYIDFNNLMAYDSSGSDKVS 249

OY 291 TYMKELYSDYAYGKKYDYIIMGFTLMFPSTPPNNDKMLVSGDFVTEKLNKRAD 350

Db 250 GMSNVFPS-----TTKPESTPFSDD-----KAVKDYIKAGVPANKIVL 288

OY 351 GFGLM 355

Db 289 GMPLY 293

RESULT 4

CH14\_TRIHA STANDARD; PRT; 423 AA.

ID CH14\_TRIHA

AC P48827;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE 42 KDA ENDOCHITINASE PRECURSOR (EC 3.2.1.14).

GN CH142.

OS Trichoderma harzianum.

OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.

OX NCBI\_TaxID=5544;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-52; 93-107; 371-385 & 397-414.

RX MEDLINE=95269313; PubMed=7750151;

RA Garcia I., Lora J.M., de la Cruz J., Benitez T., Llobell A.,

RT "Cloning and characterization of a chitinase (ch142) cDNA from the mycoparasitic fungus Trichoderma harzianum.";

RL Curr. Genet. 27:83-89(1994).

CC -1- FUNCTION: MORPHOGENETIC ROLE DURING APICAL GROWTH, CELL DIVISION AND DIFFERENTIATION (CELL WALL MORPHOGENESIS). ANTIFUNGAL AGENT.

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- INDUCTION: SPECIFICALLY INDUCED BY CHITIN AND IS CARBOLITE REPRODUCED.

CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: S78423; AAB34355.1; -

DR HSSP: P07254; ICTN.

DR InterPro: IPR001579; Chitinase\_2.

DR InterPro: IPR001223; Glyco\_hydro\_18.

DR Pfam: PF00704; Glyco\_hydro\_18; 1.

DR PROSITE: PS01095; CHITINASE\_18; 1.

KW Hydrolyase; Glycosidase; Chitin degradation; signal; zymogen;

KW Chitin-binding.

FT SIGNAL 1 22

FT PROPEP 23 34

FT CHAIN 35 423

42 KDA ENDOCHITINASE.

FT ACT\_SITE 171 171 PROTON DONOR (BY SIMILARITY).  
 FT CARBOHYD 218 218 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 423 AA; 46056 MW; DF90378BED1C30BE CRC64;

Query Match 6.6%; Score 130.5; DB 1; Length 423;  
 Best Local Similarity 21.6%; Pred. No. 0.053;  
 Matches 83; Conservative 56; Mismatches 131; Indels 115; Gaps 19;

QY 8 SFL--VICLLCLSTVSVYEGHRRARGESRKNPREIKTFKESGKIIOG-YIPSNVS 64  
 DB 3 SFLGKSVALLAALQATLSSPKRGHRRASVEKRN-----GYANSVFTMGIGI 49  
 QY 65 YNNHLK--DLNPNLVNVMHSEKMDLSYDSIESIYSGPLFKSLGLEIY-----IG 113  
 DB 50 YDRNFQPADLVAS-DYVHVHVSFPMLOAD-----GTVISQDTYADYKHNADSDMNVG 102  
 QY 114 LNEY--FNDAMNLRKAPDITMLISLGGETYHPSSFDAL-----NAVERIANLVDEL 164  
 DB 103 TNAVGCYKQLFYKVKANRGLVLLSIGGWTW-STNFPASASTDANRKNFAKTATIFMKDW 161  
 QY 165 GFGIDVDYEPNGSFDGLNDEKADFPYQVYTKLREVMCDKLISISSNGALSCIGFN 224  
 DB 162 GFGIDIDW-----ETPAD-----ATQASNNILLL----- 186  
 QY 225 DPKRICMDEAPYNSKYFNKPKVKKELRAAQAASAGAIYLMNNLKMDIMVEVQTFNY 284  
 DB 187 --KEVRSQORDA-YAAQY--AVGYHFLTLIAPAGKDNYSKRLADLGGVLYINLMAYDY 241  
 QY 285 TNSTDSIVME--LYDSIAYIKKDYIYIINGFTLMPSS-----IPFNNDKMLKSIDDF 338  
 DB 242 AGFSFPLTGHANLFNN-----PSNDAAPPFNTD-----SAVDXY 276  
 QY 339 VKTENKLNKRADGFGELSLSSDNA 363  
 DB 277 INGVPRANKIVLGMPIYGRSPQNTA 301

RESULT 5  
 CHIA\_SERMA STANDARD; PRT; 563 AA.  
 ID CHIA\_SERMA STANDARD; PRT; 563 AA.  
 AC P07254; 054275;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CHITINASE A PRECURSOR (EC 3.2.1.14).  
 GN CHIA.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 NCBI\_TaxID=615;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Koo J.C., Lim C.O., Choi Y.J., Kim C.Y., Bank J.D., Lee S.Y.,  
 RA Cho M.J.;  
 RA Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN-ATCC 990 / OMB1466;  
 RA Jones J.D.G., Grady K.L., Suslow T.V., Bedbrook J.R.;  
 RT "Isolation and characterization of genes encoding two chitinase  
 RT enzymes from Serratia marcescens.";  
 RL EMBO J. 5:467-473(1986).  
 [3]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 24-31.  
 RC STRAIN-BJL200;  
 RA MEDLINE=95154677; PubMed=7851747;  
 RA Bruberg M.B., Eljstink V.G.H., Nes I.F.;  
 RT "Characterization of a chitinase gene (chia) from Serratia marcescens  
 RT BJI200 and one-step purification of the gene product.";  
 RL FEMS Microbiol. Lett. 124:399-404(1994).  
 [4]  
 RN REVISIONS, AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).

RX MEDLINE=95219379; PubMed=7704527;  
 RA Perrakis A., Tews I., Dauter Z., Oppenheim A.B., Chet I., Wilson K.S.,  
 RA Vorgias C.E.;  
 RT "Crystal structure of a bacterial chitinase at 2.3-A resolution.";  
 RL Structure 2:1169-1180(1994).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF  
 CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L01455; AAA26551.1; -;  
 DR EMBL; X03657; CAA27292.1; -;  
 DR EMBL; Z36294; CAA85291.1; -;  
 DR PIR; A25090; A25090.  
 DR PDB; 1CTN; 20-JUL-95.  
 DR InterPro; IPR001579; Chitinase\_2.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR InterPro; IPR000601; PKD\_domain.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR SMART; SM00089; PKD; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; 1.  
 KW Hydrolyase; Glycosidase; Chitin degradation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 563  
 FT DOMAIN 150 563  
 FT ACT\_SITE 315 315  
 FT ACT\_SITE 391 391  
 FT CONFLICT 52 52  
 FT CONFLICT 73 73  
 FT CONFLICT 76 77  
 FT CONFLICT 79 79  
 FT CONFLICT 121 121  
 FT CONFLICT 139 139  
 FT CONFLICT 226 226  
 FT CONFLICT 395 395  
 FT CONFLICT 410 430  
 FT CONFLICT 437 437  
 FT CONFLICT 464 467  
 FT CONFLICT 473 473  
 FT CONFLICT 484 484  
 SQ SEQUENCE 563 AA; 60979 MW; 0696FEF6AF83AA35 CRC64;

Query Match 6.5%; Score 128.5; DB 1; Length 563;  
 Best Local Similarity 20.7%; Pred. No. 0.11;  
 Matches 97; Conservative 63; Mismatches 179; Indels 129; Gaps 24;

QY 2 NFTV-----KSYFLVYICLLC-----CLSTYVSVI-----EGHRRARGESRKNP-REIKTFK 48  
 DB 99 NFKXNKGGRGMQV--ALCNADGCTASDAIEIYVADIDGSHLP--LKEPFLKKNPKYK 153  
 QY 49 ESGGIIIGYPPSVSNHNLK-DLNPNVNVMHSEKMDLS-----YDSIESIVGS-PL 101  
 DB 154 QNSGKVVGSYFVEWGVGRNFTVDKIIPQNLFIHLVGFIPICGNGINSLKEIDGSPQA 213  
 QY 102 LFKSLIGLEYIGLNEY-----FNDAMNLRKAPDITMLISL 138  
 DB 214 LQSCQGRDEPKXSIHDPFALDQAQGVYAMDPPYKNGGQMLAKQAHPDILKIPSG 273  
 QY 139 GETYHPSSF-----DSALNAVERIANLVDELFGIDVDY-----PNSGFGINDKEX 187  
 DB 274 GWTLSDFPFPGDKVKKADRVGVSKF--LQTKFPGGVIDIDEPFGGKANPNLSPD 331  
 QY 188 ADFEVQVYTKLREYMCD-----KLISISSNGALSCIGFNDPKK-----ICM 231

```

Db 332 GETTYLMLKELAA-MDOLSAETGRKKYELTSAISGKDKIDKAVNAVQNSMDHIFLMSY 390
Qy 232 DDEAFYNSKYFN-----KPDVKKELLRAQOMASA-----GCAIY----- 265
Db 391 DEYGFDFDLKNGHOTALNAPAKKPTATYTVGVNALLAQGVKPKYVGTAMRGWTG 450
Qy 266 ---LNNNLKMDIMVYVOTFNTNSTDSTVMKELYDSTAYYK-----KDYVLIIMGFT 316
Db 451 VNGYNNIIP-----FTGATGPKVKGKTKWNGIYDROIQOGFSGEMQYTYDATAEAP 502
Qy 317 LMF-FSTP--FNPNDKMLVKSIGDFVTKENKLRKARDFGLMSLSDN 361
Db 503 YVEKPSDGLTFEDDARSVQAKGYV-----LDKOLGGLFSWEIDADN 545

RESULT 6
CHIT_STRLI STANDARD: PRT: 619 AA.
AC P36909;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE CHITINASE C PRECURSOR (EC 3.2.1.14).
GN CHIC.
OS Streptomyces lividans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1916;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=93294525; PubMed=8515228;
RA Fujii T., Miyashita K.;
RT "Multiple domain structure in a chitinase gene (chic) of Streptomyces
  lividans";
RE J. Gen. Microbiol. 139:677-686(1993).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- INDUCTION: BY CHITIN.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 BACTERIAL-TYPE CELLULOSE-BINDING DOMAIN
CC (CBD).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D12647; BAA02168.1; -
DR HSSP: P07986; IEXH.
DR InterPro: IPR001919; CBD_2.
DR InterPro: IPR001579; Chitinase_2.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001777; FN.III.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00553; CBD_2; 1.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR SMART: SMO0060; FN3; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS00561; CBD_BACTERIAL; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Hydroxylase; Glycosidase; Chitin degradation; Chitin-binding; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 619 CHITINASE C.
FT DOMAIN 35 140 CELLULOSE-BINDING.
FT DOMAIN 148 230 FIBRONECTIN TYPE-III.
FT DOMAIN 240 619 CATALYTIC.

```

```

FT ACT_SITE 382 382 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 619 AA; 65200 MW; A23CE5B3C5D6F21 CRC64;

Query Match
Best Local Similarity 20.7%; Pred. No. 0.16;
Matches 73; Conservative 54; Mismatches 105; Indels 121; Gaps 16;

Qy 50 SGRGIIGGYPSWVS--NHNLDLNPILN---VYHM-----SFAKMDL 88
Db 227 TGAEVKMGVFTWGVYGRNYHKNLYTSSADKITHINAFGNVGGKCTIDSYADYDK 296
Qy 89 STDSTESTIVG-----SPLFKSLIGLEYTGLNEFPNDAMNLRKRPDILMLSLGELY 142
Db 297 AYTAQSVQVGYADTWDP-----LRGNFNOLRLKAKYPRNKIILYSFGWTV 343
Qy 143 -----HPSPFSALNAVEKIANLVDELGFDDGIDVYE--PNSPFGDINKERADFV 192
Db 344 SGGFPAKKNPAPAFKSCDVLDEPRADY--FDGIDLDWEYFNACGLSCDTSAPNAFS 401
Qy 193 QVYTKLREYMCDDLISISOSNGALSCIGFNDPKRICMDEAPYN--SKYENKPDVYKE 250
Db 402 SMKMRMRAEFQGDYILTAVTADG-----SDGKTL--DAADYGEASKY----- 442
Qy 251 LIRAAQMASAGAIYLMNNLKDMDVYVOTFNTNSTDSTVMKELYDSTAYYK----- 304
Db 443 -----IDMYNMYTFD-----FAGMANGPTAPHS 467
Qy 305 --KRYDYVIMGFTLMPSTPEN---PNDKMLVKSIGDFVTKENKLRKRDG 351
Db 468 PLTAYDIGIPQGFNTADAMAKKSKGVPAKLLI-GIGFYGNGWTVQVSAPG 519

RESULT 7
CHIT_BRUMA STANDARD: PRT: 504 AA.
AC P29030;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE ENDOCHITINASE PRECURSOR (EC 3.2.1.14) (MFI ANTIGEN).
OS Brugia malayi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N. A., AND PARTIAL SEQUENCE.
RX MEDLINE=92179220; PubMed=1542646;
RA Fuhman J.A., Lane W.S., Smith R.F., Plessens W.F., Perlier F.B.;
RT "Translational-blocking antibodies recognize microfilarial chitinase
  in brugian lymphatic filariasis.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1548-1552(1992).
CC -1- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY
CC FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-
CC FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND
CC TRANSMISSION.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND
CC WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.
CC -1- PTM: O-GLYCOSYLATED.
CC -1- MISCELLANEOUS: KNOWN TO BIND CALCIUM.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```







Dd		320	VIAAGSGSVYTKQVIGTRELIFRANFPPSHAQTIGFTQLSDNCGWGDMDWTGNNGNE	379
Oy	328	DKMLV----	KSIQGVKTEKN-LNRADGFGSLSSDNA-----HNEOLAIEY	372
Dd		380	IYMRVGYSGGSDIGDVPNSNKYLK----FYIYDLKRISTTDLNFETIDLDIEY	431
RESULT	11			
ID	DRTS_PLABA	STANDARD:	PRT:	587 AA.
AC	Q27713; Q27714;			
DT	01-NOV-1997 (Rel. 35, last sequence update)			
DT	01-NOV-1997 (Rel. 35, last annotation update)			
DT	15-JUL-1999 (Rel. 36, last annotation update)			
DE	BIOCHEMICAL DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (DHFR-TS)			
DE	(INCLUDES: DIHYDROFOLATE REDUCTASE (EC 1.5.1.3); THYMIDYLATE SYNTHASE (EC 2.1.1.45)).			
DS	Plasmodium berghei (strain Anka).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5833;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95198769; PubMed=7891743;			
RA	van Dijk M.R., McConekey G.A., Vinkenoog R., Waters A.P., Janse C.J.;			
RT	"Mechanisms of pyrimethamine resistance in two different strains of Plasmodium berghei."			
RL	Mol. Biochem. Parasitol. 68:167-171(1994).			
RN	[2]			
RP	SEQUENCE OF 12-201 FROM N.A.			
RX	MEDLINE=95059225; Pubmed=7969277;			
RA	Cheng Q., Saul A.;			
RT	"The dihydrofolate reductase domain of rodent malarial point mutations and pyrimethamine resistance."			
RL	Mol. Biochem. Parasitol. 65:361-363(1994).			
CC	-1- CATALYTIC ACTIVITY: 5,6,7,8-TETRAHYDROFOLATE + NADP(+) =			
CC	7,8-DIHYDROFOLATE + NADPH.			
CC	-1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =			
CC	DIHYDROFOLATE + DTMP.			
CC	-1- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS,			
CC	DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO DTMP.			
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE			
CC	DIHYDROFOLATE REDUCTASE FAMILY.			
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE			
CC	SYNTASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/isb-sib.ch">http://www.isb-sib.ch/announce/isb-sib.ch</a> ).			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U12275; AAB60237.1; -			
DR	EMBL; L28119; AAA29581.1; -			
DR	HSSP; P00470; IANS.			
DR	InterPro; IPR001796; DHFR.			
DR	InterPro; IPR000398; Thymidylat_synt.			
DR	Pfam; PF00186; Dihfolate_red. 1.			
DR	Pfam; PF00303; thymidylat_synt. 1.			
DR	PRINTS; PR00070; DHFR.			
DR	PRINTS; PR00108; THYMSNTASE.			
DR	ProdDom; PD001180; Thymidylat_synt. 1.			
DR	PROSITE; PS00075; DHFR. 1.			
DR	PROSITE; PS00091; THYMIDYLATE SYNTHASE. 1.			
DR	Multifunctional enzyme; Oxidoreductase; 1.			
DM	Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.			
FM	DOMAIN 1 240			
FT	DOMAIN 301 587			
FT	ACT_SITE 469 469			
FT	VARIANT 110 110			
FT	VARIANT 177 177			
FT	S -> F (IN PYRIMETHAMINE RESISTANCE).			
FT	S -> F (IN PYRIMETHAMINE RESISTANCE).			

SEQ	SEQUENCE	587 AA;	68932 MW;	6E638C2B02FEC13A	CRC64;
QY	7 YSFVILCLLCCLSTFYVSVEGHRARPGESRKNPBEIIITKFEESGK-----IIQG 57	5.4%;	Score 108;	DB 1;	Length 587;
	Best Local Similarity 18.7%; Pred. No. 2.9;				
	Matches 80; Conservative 67; Mismatches 132; Indels 148; Gaps 20;				
DB	8 FETIVACACCKVLN-----DDEKVCRFNN-----KTEGIGNAGVLPKCNLIDMK 53				
QY	58 YPSPWVS-----NHNLK-----DLNPUL-----NVVM-----SEAK 85				
DB	54 YSSVTSYINENNNYIRLKKRKDKYMEKHNLKNVLENTNIISTNNLQNIYVGRKKSMS 113				
QY	86 MDLSIDSEISVGPLKSLGLEIYIGINEFNDAMLRKRPDIIMLS-----136				
DB	114 IPKKFPLQNRNIILSLTKKEEDIVNNNNNNVYIIKSVDDLFLPLCKTKYKCFI 172				
QY	137 LGGEFYHPSPFSALNAVAKIANLVDELGFSDIDVDEPNSFGDGNDEKADFVQYVT 196				
DB	173 IGGSSYKFFLD-----KLLIKITFTIINNSYNDVLPPELNE-----NLF-----214				
QY	197 KLEFYMCDELKILISQ---SSNGALSCIGFNDPKKICMDDEAPYNS-----239				
DB	215 -----KINSISDVYVSNNTTIDFLITYSKTEINPEEVPNNFTLGCVEQNKAFDD 265				
QY	240 ---KYF---NKPDYKELLPAQMASGAGIYLMNNIKMDIADVQTFYITSTST 291				
DB	266 EDDYTYFSEFNKKENIKKNSEH-----HNKR-----IYNSIKYKNNPEYO 306				
QY	292 VKLELYDSVAYYKKYDYI-----MGFTLMFSPSTFNP---NDKMLVYSIGD---FV 339				
DB	307 YNIITIDIMHGNKQDDRIGVGLSKFYIMKFNLINEYPLLTTKLPIRGIIEELMPTI 366				
QY	340 KTEENKLN 346				
DB	367 RGETNGN 373				
RESULT 12	CHIT MANSE	STANDARD:	PRT:	554 AA.	
AC	P36362:				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DT	01-FEB-1995 (Rel. 31, Last annotation update)				
DE	ENDOCHITINASE PRECURSOR (EC 3.2.1.14).				
OS	Manduca sexta (tobacco hawkmoth) (Tobacco hornworm).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;				
OC	Sphingioidea; Sphingidae; Sphinginae; Manduca.				
OX	NCBI_TaxID=7130;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93357793; PubMed=8353525; Muthukrishnan S.;				
RA	Kramer K.J., Corpuz L., Choi H.K.,				
RA	"Sequence of a cDNA and expression of the gene encoding epidermal and				
RT	gut chitinases of Manduca sexta";				
RL	Insect Biochem. Mol. Biol. 23:691-701(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=97215580; PubMed=9061927;				
RA	Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;				
RA	"Isolation and characterization of a genomic clone for the gene of an				
RT	insect molting enzyme, chitinase";				
RL	Insect Biochem. Mol. Biol. 27:37-47(1997).				
CC	-1- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING				
CC	PROCESS.				
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF				
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.				
CC	-1- SUBCELLULAR LOCATION: SECRETED.				
CC	-1- TISSUE SPECIFICITY: EPIDERMIS AND GUT.				









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 08:57:32 ; Search time 26.17 Seconds  
(without alignments)  
2112.760 Million cell updates/sec

Title: US-09-579-383-3

Perfect score: 1982  
Sequence: 1 MNFTVKYSFLVICLLCCLLS.....SDNAANEQLAIEYFVESLH 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP:REMBL\_17.\*
- 2: SP:archaea.\*
- 3: SP:bacteria.\*
- 4: SP:fungi.\*
- 5: SP:human.\*
- 6: SP:invertebrate.\*
- 7: SP:mhc.\*
- 8: SP:organelle.\*
- 9: SP:phage.\*
- 10: SP:plant.\*
- 11: SP:rodent.\*
- 12: SP:virus.\*
- 13: SP:vertebrate.\*
- 14: SP:unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1982	100.0	378	5 Q9U419	Q9U419 plasmodium
2	1694	85.5	323	5 Q9NUX8	Q9NUX8 plasmodium
3	635	32.0	587	5 Q9NUX9	Q9NUX9 plasmodium
4	175.5	8.9	699	2 Q48494	Q48494 kurtzia zop
5	167.5	8.5	832	2 Q92NLO	Q92NLO clostridium
6	160.5	8.1	717	2 Q9KHB3	Q9KHB3 bacillus ci
7	156.5	7.9	405	5 Q97403	Q97403 phaeodon coc
8	149	7.5	635	2 Q07088	Q07088 bacillus th
9	148.5	7.5	496	2 Q9RG51	Q9RG51 bacillus ci
10	146.5	7.4	674	2 Q9FAC8	Q9FAC8 bacillus ce
11	146.5	7.4	1088	2 Q9KT32	Q9KT32 vibrio chol
12	146	7.4	560	3 Q9P401	Q9P401 ejellomyces
13	142	7.2	493	3 Q9HEW6	Q9HEW6 ejellomyces
14	142	7.2	604	3 P94339	P94339 bacillus li
15	139	7.0	729	2 P96168	P96168 vibrio harv
16	138.5	6.9	563	2 Q9ALZ0	Q9ALZ0 serratia li
17	137.5	6.9	427	3 Q00435	Q00435 coccidioid
18	136.5	6.9	562	2 P97034	P97034 enterobacte
19	133.5	6.7	416	3 Q92222	Q92222 emericella

20	132.5	6.7	553	5 P91731	P91731 hyphantria
21	132.5	6.7	563	2 O54328	O54328 enterobacte
22	132.5	6.7	599	2 Q9KED7	Q9KED7 bacillus ha
23	130.5	6.6	491	2 P94289	P94289 bacillus ci
24	130	6.6	443	3 Q9HG05	Q9HG05 coniothyriu
25	129.5	6.5	424	3 Q9Y841	Q9Y841 metarhizium
26	128.5	6.5	286	10 Q9YFR9	Q9YFR9 oryza sativ
27	127.5	6.4	404	3 Q9P400	Q9P400 ejellomyces
28	126.5	6.4	609	2 Q9Z9M8	Q9Z9M8 streptomyce
29	126.5	6.4	1054	2 Q9ZIX4	Q9ZIX4 pseudodalter
30	126	6.4	400	3 Q92270	Q92270 rhizopus ol
31	126	6.4	411	3 Q9UV48	Q9UV48 trichoderma
32	126	6.4	555	5 Q9U6R7	Q9U6R7 dermatophy
33	125.5	6.3	543	5 Q9GV05	Q9GV05 bombyx mori
34	125.5	6.3	543	5 Q9GR93	Q9GR93 bombyx mori
35	125.5	6.3	563	2 Q83008	Q83008 serratia ma
36	125.5	6.3	565	5 P90710	P90710 bombyx mori
37	125.5	6.3	594	12 Q9PYU0	Q9PYU0 xestia c-ni
38	125	6.3	596	2 Q85500	Q85500 bacillus su
39	124.5	6.3	511	3 Q06350	Q06350 saccharomyc
40	124	6.3	418	3 Q9UV45	Q9UV45 hypocreata ru
41	123.5	6.2	417	2 Q9Z9M7	Q9Z9M7 streptomyce
42	122.5	6.2	412	3 Q9UV40	Q9UV40 trichoderma
43	122.5	6.2	696	5 Q9VZV2	Q9VZV2 dirosophila
44	122	6.2	360	2 Q9FAC9	Q9FAC9 bacillus ce
45	122	6.2	424	3 Q9CIT9	Q9CIT9 trichoderma

ALIGNMENTS

RESULT 1  
ID Q9U419 PRELIMINARY: PRT: 378 AA.  
AC Q9U419:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CHITINASE.  
OS Plasmodium falciptarum (Isolate KI / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KI;  
RX MEDLINE=20040676; PubMed=10570198;  
RA Vinetz J.M., Dave S.K., Specht C.A., Brameld K.A., Kuntz I.D., Xu B.,  
RT Hayward R., Fidock D.A.;  
RT "The chitinase PICH1 from the human malaria parasite Plasmodium  
RT falciptarum lacks proenzyme and chitin-binding domains and displays  
RT unique substrate preferences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:14061-14066(1999).  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL: AF172445; AAF16902.1; -;  
DR InterPro: IPR001579; Chitinase\_2.  
DR InterPro: IPR002114; PTS\_HPR\_ser.  
DR PROSITE: PS01095; CHITINASE\_18; 1.  
DR PROSITE: PS00589; PTS\_HPR\_SER; UNKNOWN\_1.  
KW Glycosidase; Hydrolase.  
SQ SEQUENCE 378 AA: 42786 MW; 5DF86F2AC93C8IE3 CRC64;

Query Match 100.0%; Score 1982; DB 5; Length 378;  
Best Local Similarity 100.0%; Pred. No. 1.9e-131;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNFTVKYSFLVICLLCCLSTVSVYIEGHRAPGSRKNPREITKFKSGGIIGYPP 60  
DB 1 MNFTVKYSFLVICLLCCLSTVSVYIEGHRAPGSRKNPREITKFKSGGIIGYPP 60  
QY 61 SVVSTNNHNLKDLNPNLVNHYHMSFAKMDLSYDSIESIVGSPLLFKSLIGETIGLNEYFND 120  
|||||

```
Db 61 SWVSYNHNLKDLNPNLVNVMHSAKMDLSYDSIESIVGSPLLFSLGLEIYGLNMFND 120
Qy 121 AMNLKRPADIIIMLSLIGETYPSPSALNAVEKIANLVDELGPDIDVYEPNGSD 180
Db 121 AMNLKRPADIIIMLSLIGETYPSPSALNAVEKIANLVDELGPDIDVYEPNGSD 180
Qy 181 GLNDEKADFEVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMDDEAPYNSK 240
Db 181 GLNDEKADFEVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICMDDEAPYNSK 240
Qy 241 YFNKPDVAKELLRAQOMASAGAIYLMNKLKMDIMVFQVTNTNSTSTYWKELYSY 300
Db 241 YFNKPDVAKELLRAQOMASAGAIYLMNKLKMDIMVFQVTNTNSTSTYWKELYSY 300
Qy 301 AYGKRYVYIIMGFTLMEPTSPFNPKMLVKSIGDFYTKENKLRADGFLMSLSSD 360
Db 301 AYGKRYVYIIMGFTLMEPTSPFNPKMLVKSIGDFYTKENKLRADGFLMSLSSD 360
Qy 361 NAAHNEOLAIEYVESLH 378
Db 361 NAAHNEOLAIEYVESLH 378

RESULT 2
Q9NXX8 PRELIMINARY; PRT; 323 AA.
AC Q9NXX8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHITINASE (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20209408; PubMed=10744721;
RA Vinetz J.M., Valenzuela J.G., Specht C.A., Aravind L., Langer R.C.,
  Ribeiro J.M.C., Kaslow D.C.;
RT "Chitinases of the Avian Malaria Parasite Plasmodium gallinaceum, a
  Class of Enzymes Necessary for Parasite Invasion of the Mosquito
  Midgut."
RT J. Biol. Chem. 275:10331-10341(2000).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
  HYDROLASES).
DR EMBL: AF072442; AAF63209.1; -
  InterPro: IPR001579; Chitinase.2.
DR InterPro: IPR002114; Pts_HPR_ser.
DR PROSITE: PS01095; CHITINASE_18; 1.
DR PROSITE: PS00589; Pts_HPR_SER; UNKNOWN.1.
KW Glycosidase; Hydrolase.
FT NON_TER 1 1
FT NON_TER 323 323
SQ SEQUENCE 323 AA; 36502 MM; 9079ADEFE0F754D2 CRC64;
```

Query Match 85.5%; Score 1694; DB 5; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.5e-111;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 52 KGIIGYPPSVSYNNHNLKDLNPNLVNVMHSAKMDLSYDSIESIVGSPLLFSLGLEIY 111
Db 1 KGIIGYPPSVSYNNHNLKDLNPNLVNVMHSAKMDLSYDSIESIVGSPLLFSLGLEIY 60
Qy 112 IGLNEYNDAMNLKRPADIIIMLSLIGETYPSPSALNAVEKIANLVDELGPDIDV 171
Db 61 IGLNEYNDAMNLKRPADIIIMLSLIGETYPSPSALNAVEKIANLVDELGPDIDV 120
Qy 172 DYENSGFDGLNDEKADFEVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICM 231
Db 121 DYENSGFDGLNDEKADFEVQYVTKLREYMCDDKLISISQSSNGALSCIGFNDPKKICM 180
Qy 232 DDEAPYNSKYFNKPDVAKELLRAQOMASAGAIYLMNKLKMDIMVFQVTNTNSTST 291
```

```
Db 181 DDEAPYNSKYFNKPDVAKELLRAQOMASAGAIYLMNKLKMDIMVFQVTNTNSTST 240
Qy 292 VKKELYDAAAYGKRYDVYIIMGFTLMEPTSPFNPKMLVKSIGDFYTKENKLRADG 351
Db 241 VKKELYDAAAYGKRYDVYIIMGFTLMEPTSPFNPKMLVKSIGDFYTKENKLRADG 300
Qy 352 FGLMSLSDNAAHNEOLAIEYFV 374
Db 301 FGLMSLSDNAAHNEOLAIEYFV 323
```

```
RESULT 3
Q9NXX9 PRELIMINARY; PRT; 587 AA.
AC Q9NXX9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENDOCHITINASE PRECURSOR (EC 3.2.1.14).
OS Plasmodium gallinaceum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5849;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20209408; PubMed=10744721;
RA Vinetz J.M., Valenzuela J.G., Specht C.A., Aravind L., Langer R.C.,
  Ribeiro J.M.C., Kaslow D.C.;
RT "Chitinases of the Avian Malaria Parasite Plasmodium gallinaceum, a
  Class of Enzymes Necessary for Parasite Invasion of the Mosquito
  Midgut."
RT J. Biol. Chem. 275:10331-10341(2000).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
  HYDROLASES).
DR EMBL: AF064079; AAF63208.1; -
  InterPro: IPR001579; Chitinase.2.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 587 AA; 67918 MM; 23D1ABA87DEBEE37 CRC64;
```

Query Match 32.0%; Score 635; DB 5; Length 587;  
Best Local Similarity 37.4%; Pred. No. 1.4e-36;  
Matches 129; Conservative 82; Mismatches 126; Indels 8; Gaps 5;

```
Qy 37 RKNPREIITKFEKSGKGIIGYPPSVSYNNHNLK---DLNPNLVNVMHSAKMDLSYDSI 93
Db 89 RKNPQIIEYKRRKQGIAGYGSNMSGRAKHMDSNPMVSLIYAFARIMLYDVS 148
Qy 94 ESIYVSPLLFSLGLEIYIGNEYFNDAAMLKRPADIIIMLSLIGETYPSPSALNA 153
Db 149 RPNRQRFLLRKHGLEYETYGMMLEIRIRKVPDVIILSLGGERY-MIDIEKIDY 207
Qy 154 VEKIANLVDELGPDIDVYEPNGSFDGLNDEKADFEVQYVTKLREYMCDDKLISISQ 213
Db 208 VDKILKLVNDFDLGDVIDMEPHGKFNELNLSNYIKILINLRTIPEKILISISG 267
Qy 214 SNGALSCIGFNDPKKICMDDEAPYNSKYFNK-PVKKELLRAQOMASAGAIYLMNKLK 272
Db 268 SNAALSCY--SGVASFCDESPYNTKFLSQIETKELHRAAAMLSAGPFINFNPAKE 325
Qy 273 MIDVVFQVTNTNSTSTYWKELYSYAYYGGKRYDVYIIMGFTLMEPTSPFNPKMLV 332
Db 326 KIDLVFIQTYN-LETTNDINVDWYLSHLVGLKYNITIIILGFLHNRRGFSPEKNEL 384
Qy 333 KSIQDFYTKENKLRADGFLMSLSDNAAHNEOLAIEYVESL 377
Db 385 ELVGKTIHDKQNNNRADGIGIWHLFMKEQLPFGSFDVDFLFTNI 429
```

```
RESULT 4
Q48494 PRELIMINARY; PRT; 699 AA.
```



AC 048494;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE CHITINASE.  
 OS Kurthia zopfi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Kurthia.  
 OX NCBI\_TaxID=1650;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ikeda S., Toyoda H., Matsuda Y., Ouchi S.;  
 RT "DNA sequence determination of a chitinase gene chisH1 cloned from  
 RT gram-positive bacterium Kurthia zopfi and its application to  
 RT biological control of Powdery Mildew of Barley.";  
 RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 DR EMBL: D63702; BAA09831.1; -.  
 DR HSSP: P07254; ICTN.  
 DR InterPro: IPR001579; Chitinase-2.  
 DR InterPro: IPR003610; Chitin\_bind3.  
 DR InterPro: IPR003962; FNIII\_repeat.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00041; fn3; 2.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR PRINTS: PR00014; ENTPYELII.  
 DR SMART: SM00495; ChIBD3; 1.  
 DR SMART: SM00060; FN3; 2.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR Glycosidase; Hydrolase; Repeat.  
 KW SEQUENCE 699 AA; 73494 MW; 2AE959A604BF513 CRC64;

Query Match 8.9%; Score 175.5; DB 2; Length 699;  
 Best Local Similarity 20.1%; Pred. No. 0.00034;  
 Matches 89; Conservative 59; Mismatches 109; Indels 185; Gaps 22;

55 IGGYPSWVS--NHNLTGKLNPLNVHMSFAKMDLSYDSIES----- 95  
 46 IGGYPSWAGYGRNVWDIDPT-KVTHINVAADICMGVHGNDPPSGPNVPTWCONE 104  
 96 -----IVGSPILFKSLIGLEYIG-----LNEYFNDAMLRKARPDIIIMLSL 137  
 105 KSQTIWVPGTIVLGDPTWID---TGKTFAGDPTWIDP IAGINIQLNKLRKINERLKITISV 161  
 138 GGETYHPSPFSALNAV--EKIAN-LVDEL---GPDGIDVYE--PNSGPDGLNDEKAD 189  
 162 GGMTWNSRFSDDVAATPAATREVFANSVDFLRKYNEPDGVDLWEYPSGGLDG-NSKRSAD 220  
 190 --FEVQYVTKLEBYMDDKLISISOSNGALSICGFNDPKKICMDEAIVNYSKYFKRPV 247  
 221 KONYTILLSKIREKL-----DAAGAV-----DG 243  
 248 KELLRAAOMASAGAIYLMN---NLKMDIMVYQVFNTN---STDSTVMKELYD 298  
 244 KYYLL--TIASGASTYVAANTELANIASIVDMINIMTDFGANGKQSAHANALLN--YD 298  
 299 SYA-----YGGKKYD----- 308  
 299 PAASAGVDPANTFVNAAGAGHLNAGVPAKLVLGVPYGGMGCAQAGNGOYOTCTG 358  
 309 -----YIIMGFTLMPSTP-----FNPNDKMLV-----KSIQDF 338  
 359 GSSVGTWEGSDFDYDLEINYNKNGCYTRYWMDTAKVPIYVNASNKRFTSYDAESIG-- 416  
 339 VKTENKLNKRRADGFGIMLSLSD 360  
 417 YKTAYIKSKGLGAMFWELSGD 438

RESULT 5  
 ID Q92N10 PRELIMINARY; PRT; 832 AA.  
 AC Q92N10;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE CHITINASE A.  
 GN CHIA.  
 OS Clostridium paraputrificum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 CC Clostridium.  
 OX NCBI\_TaxID=29363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-M;  
 RA Morimoto K., Karita S., Kimura T., Sakka K., Ohmiya K.;  
 RT "Cloning, Sequencing, and Expression of the Gene Encoding Clostridium  
 RT Paraputrificum Chitinase ChIA and Analysis of the ChI Transcription.";  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 DR EMBL: AB012764; BAA34922.1; -.  
 DR InterPro: IPR001579; Chitinase-2.  
 DR InterPro: IPR003610; Chitin\_bind3.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 3.  
 DR SMART: SM00495; ChIBD3; 1.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 832 AA; 92586 MW; BBB6FE9B04370478 CRC64;

Query Match 8.5%; Score 167.5; DB 2; Length 832;  
 Best Local Similarity 24.9%; Pred. No. 0.0016;  
 Matches 90; Conservative 50; Mismatches 130; Indels 91; Gaps 19;

9 FVTCILLCILSTVYSVIEGHRARPEGSRKNPREIITKPKESG--KGIIGYPSWVSYN 66  
 9 FLSTLLVVALTSAGITVTSAR---EAKDNGQLVNTNKNDELKRIKIVGPEP-AVS 63  
 67 -----HNKDLN-PLNVNVHMSFAKMDLSYDSI-----ESTVSPILFKSLT-- 107  
 64 SEAGQFFNATDQLQMSLTHIQYSFAVVDPTNKTITLGDKHAIEEDPFTGYDLSYKGIYE 123  
 108 -----GLEYGILNEYFNDAMLRKARPDIIIMLSIGFT-----YHPSPSALNA-VEKIA 158  
 124 LDTSLPYKG--HFNVLTQTKKSYPDVLLISVGGAGSRGFTYMLDPTAGINTFPADSCV 180  
 159 NLVDELGPDGIDVDEPEPNSGPDGLNDEKADFEVQYVTKLREYMCDDKLISISOSNGAL 218  
 181 EPIRQYGFQGVVIDEYEPATSQSGNPDPDFLSEPRSRKLNRY--NLIKITLR----- 232  
 219 SCITGENDPKKICMDEAPNYSKYFKPDKYKELLRAAOMAS---AGAIYILMNKLKMD 275  
 233 -----OKI--DEAA-----KADCKDYLLAAVYASPVVLGV--SDNSYAKYLD 272  
 276 WVEVOTFN-----YTNSDSTVMKELYD-----SVAYY-GKKYDVITM 313  
 273 FLVSMSYDIHGGMNEVEHLAGIYRNAEDRETYVAOIMPLCDMAVRIYRGVLPSEKILM 332  
 314 G 314  
 333 G 333  
 RESULT 6  
 ID Q9KHB3 PRELIMINARY; PRT; 717 AA.  
 AC Q9KHB3;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE CHITINASE.  
 GN CH11.  
 OS Bacillus circulans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1397;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-2;  
 RA Wang Y.L., Wang H.Y., Qin M., Zhang Y.Z.;  
 RT "Nucleotide sequence of a chitinase gene (ch11) from Bacillus  
 circulans C-2."  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF265220; AAF74782.1;  
 DR InterPro: IPR001579; Chitinase\_2.  
 DR InterPro: IPR003610; Chitin\_bind3.  
 DR InterPro: IPR003962; FnIII\_repeat.  
 DR InterPro: IPR003961; Fn\_III.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00041; fn3; 2.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR PRINTS: PR00014; FMYPEPIL.  
 DR SMART: SM00495; CHBD3; 1.  
 DR SMART: SM00060; FN3; 2.  
 DR PROSITE: PS01095; CHITINASE\_18; UNKNOWN\_1.  
 DR Hydrolase; Repeat.  
 KW SEQUENCE 717 AA; 76118 MW; BB0B019CDE72C198 CRC64;

Query Match 8.1%; Score 160.5; DB 2; Length 717;  
 Best Local Similarity 22.8%; Pred. No. 0.004;  
 Matches 105; Conservative 65; Mismatches 143; Indels 147; Gaps 27;

QY 11 VILLCCLLTSTVYS-VIEGHRARPGESRKNPRLITFKESGKIIGYPPSWSY--N 66  
 DB 16 VILGLVLLSTIPSTFQSTRTAEDAYK-----IVGYPMAAAYGRN 59  
 QY 67 HNLKDLNPNLVNVMSPAKMDLSYDSIES-----IV 97  
 DB 60 YNTVDIDPT-KVTHINAFADICNGHIGNPDSPNPVTWSCQNEKSGQINVPGLIVL 118  
 QY 98 GSPLL--FKSLIG----LEYGLNEYENDAMNLRKRPDIIMLSLGETYHPSEDSAL 151  
 DB 119 GDMWIDGKQFAGDTWQDPAG--NINQLKRLKOVNPLTLIISIGGWTSMNRESDAA 175  
 QY 152 NAV--EKIAN-LYDEL---GPDGIDV--DYEPNGSPDLINKKAD--FFQVQYTKLREY 201  
 DB 176 TSAITREVFANSADVFLKRYNFDGVDLNMWEPVSGLDG--NSKRPEDKONTYLLSKIREK 234  
 QY 202 M-----CDDK--LISISQ-----SSNGALSCIG-----FNDPKKICMDEAP 236  
 DB 235 LDAEAENVGKEYLITISGASPTAANTLAINISYDVIMIMYDFGCAQNKISAHNAP 294  
 QY 237 YN-----SKYFNKPDVKKELLRAAQMASAGAIYLMNNL---KMIDMVF--- 278  
 DB 295 LNAADPAASAGVDPDSNFTN-----VAAGAGHLNAGVPAKLVLGVPFYGR 340  
 QY 279 -----VQTFN---YTNGTSTVYMKELDYSAVYGGKYDYVINGFTLMFPT---PF--N 325  
 DB 341 GMDGCAQADNGOYOTCGGGSIGTWEAGSFDYDEANVYINKNGTYRWMDTAKVPLYN 400  
 QY 326 PNDKMLV-----KSIGDFVTEKNLNRADFGILMSLSD 360  
 DB 401 ASNKRFTSYDDAESIGH--KTAATYKSGKLGAGMFWELSGD 438

RESULT 7  
 ID 097403 PRELIMINARY; PRT; 405 AA.  
 AC 097403;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE CHITINASE PRECURSOR (EC 3.2.1.14).  
 OS Phaeodon cochleariae (Mustard beetle).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Phyltophaga; Chrysomeloidea; Chrysomellidae;  
 OC Chrysomellinae; Phaeodon.  
 OX NCBI\_TaxID=80249;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=GUT;  
 RA Girard C., Jouanin L.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 DR EMBL: Y18011; CAAT7014.1;  
 DR HSSP; P07254; ICTN.  
 DR InterPro: IPR001579; Chitinase\_2.  
 DR InterPro: IPR001064; Crystallin.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
 DR Glycosidase; Hydrolase; Signal.  
 KW SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 405 CHITINASE.  
 FT SEQUENCE 405 AA; 44995 MW; 11E900AFA66272AD CRC64;

Query Match 7.9%; Score 156.5; DB 5; Length 405;  
 Best Local Similarity 20.2%; Pred. No. 0.0033;  
 Matches 74; Conservative 63; Mismatches 114; Indels 115; Gaps 15;

QY 9 FLVILLCCLLTSTVYSVIEGHRARPGESRKNPRLITFKESGKIIGYPPSWSY--N 66  
 DB 6 FISFVLTLIRSS-ISTV-----SGRNY-CIFASWYIRPG 40  
 QY 67 --HNLKDLNPNLVNVMSPAKMDLSYDSIESIYGSPLFKSLIGLEYIGLNEYENDAMN 123  
 DB 41 NGLFVSNIEPDL-CTHINFAFGLHEDGTINIIDK--WESDDGKYHG--FRLLD 92  
 QY 124 LKRAPDIIMLSLIG-----ETHPSEDSAL--NAVEKIANLYDELGPDGIDVYEPN 176  
 DB 93 LRNSHPSLKYLVSMGNEGTRKNSKYAADVLRKTLANNVGAFTROYGDFGDIOMEYR 152  
 QY 177 GSPDLNDEKADPFVQVTKLREYM--CDDKLISOSSNGALSCIGFNDPKKICMDEA 235  
 DB 153 GSREGSNVTTIKDNFVALLIEDLSAVLHPKGLLRAAAGVERIDGFDVPR----- 204  
 QY 236 PYNKRYFNKPDVKKELLRAAQMASAGAIYLMNNLKMID-MVF----- 278  
 DB 205 -----VNEILDMINVMYDFGHREPFVGHLSPL 233  
 QY 279 -VQTFNTNSTDSTVYMKELDYSAVYGGKYDYVING-----FTLMFSTPNPNDKA 330  
 DB 234 HASSIDYENGNNATVATGIRKIWIYGASPEKINMGAIATYGRSFTLKDN-----NTDL 288  
 QY 331 LVKSIG 336  
 DB 289 YAPNAG 294

RESULT 8  
 ID 007088 PRELIMINARY; PRT; 635 AA.  
 AC 007088;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE CHITINASE.  
 GN CH1.  
 OS Bacillus thuringiensis.  
 OS Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1428;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Sonngay S., Panhangred W.;  
 RA Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 DR EMBL: U89796; AAB58579.1; -.  
 DR InterPro: IPR001919; CBD\_2.  
 DR InterPro: IPR001579; Chitinase\_2.  
 DR InterPro: IPR003861; FN\_III.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00553; CBD\_2; 1.  
 DR Pfam: PF00041; fn3; 1.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 2.  
 DR SMART: SM00060; FN3; 1.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; UNKNOWN\_1.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 635 AA; 71096 MW; 72D970E9DB1CC3AD CRC64;

Query Match 7.5%; Score 149; DB 2; Length 635;  
 Best Local Similarity 18.7%; Pred. No. 0.021;  
 Matches 85; Conservative 76; Mismatches 127; Indels 166; Gaps 19;

OY 6 KYSFVLCCLCLTCL--LSYVSVIEGHRARPGESRKNPREIKTFKESGKIIGYPSW 62  
 DB 7 KFTLLLSLILFLPLFTNFTPNLALDSPKOSK-----IVGYFPSW 50  
 OY 63 VSY--NHNKLDLPMNLVHMSFAKMDLSYDSIESIVSP-----LIFKSLI--- 107  
 DB 51 GYGNANYOVDIDAS-KLTLNLNYAFADICWNGKH--GNPFSSNQINKLDCRDLVYHC 106  
 OY 108 -----GLEYGLENE-----YENDANLKRKAPDITMLISL 137  
 DB 107 KIKVNGEPLVLENGHGLMIPNRILAQGTWEDCDYKARGNFGELRKAKYFHELTISV 166  
 OY 138 GGETYHPSFSDALNA-----VEKIANLVDELGFIDVDE-----PNSGFDGLND 184  
 DB 167 GGMTSARNSPDMADKTRKYPFAESTVAFPLPAYGFDVLDMEYGVETIEGST---RP 223  
 OY 185 KEKADFVQVYTKLREYVCDKLISISQSSNGALSCIGFNDPKKICMDDEAPYNSYENK 244  
 DB 224 EDKQKFTL-----LLOELRNALNKAGAECDCKOYLLTIPSGASQRYADH 266  
 OY 245 PDVKK-----ELRAQMASNGAIYLMNNLKMDIADVPQTFN----- 283  
 DB 267 TELKRIQSYSIGLISHMISTADGKLL-----TIMOLYIGSNDPANTNFEYDGAINV 320  
 OY 284 YTNSTDS-----TVMKELDYSAVYKKYDYVIMGFTLMPSTPEPNCKMLVK 333  
 DB 321 YTNREVSILPTQCPKRLDLVKEILGG-----PLSQKTEPT 357  
 OY 334 SIGDFVTKENLKRKADGFGI--WLSISSDMAANE 366  
 DB 358 NVKNIVT-----NKNNSNVQLNMTVSTDNVGYTE 387

RESULT 9  
 O9RG51 PRELIMINARY; PRT; 496 AA.  
 AC O9RG51;  
 DT 01-MAY-2000 (Tremblrel, 13, Created)  
 DT 01-MAY-2000 (Tremblrel, 13, last sequence update)  
 DT 01-JUN-2001 (Tremblrel, 17, last annotation update)  
 DE CHITINASE.  
 GN CH141.  
 OS Bacillus circulans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1397;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NO. 4.1;  
 RA Thepouyorn A., Wiat C.;  
 RT "Chitinase-encoding gene from Bacillus circulans";  
 CC Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RL -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 DR EMBL: AF154827; AAF23368.1; -.  
 DR HSSP: P07254; ICTN.  
 DR InterPro: IPR001579; Chitinase\_2.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 496 AA; 55969 MW; 25E7D2AF0F213649 CRC64;

Query Match 7.5%; Score 148.5; DB 2; Length 496;  
 Best Local Similarity 21.5%; Pred. No. 0.016;  
 Matches 105; Conservative 72; Mismatches 138; Indels 173; Gaps 28;

OY 7 YSFLVLCCLCLSTVYVIEGHRARPGESRKNPREIKTFKESGKI-IGYPSWVS 65  
 DB 15 FSFIFVMLLS-----LSFVNGEVAK-----ADSGKNYKIIIGYPSWAV 53  
 OY 66 NHNLK--DLNPLNVHMSFAKMDLSYS-----IES----- 95  
 DB 54 GRNFYWMMDVS-KVSHINYAFADICWEGRGHNPPTGPNPOTWSCODENGYIDAPNGTI 112  
 OY 96 IVGSPDL--FKSLIG--LEYIGLEYNFNDANLKRKAPDITMLISGGETYHPSFSDA 150  
 DB 113 VMGDWDIAQNSPNPDVDEPIRGN--FKOLLKTKSHPLKTFISVGWTFWSNRFSDVA 170  
 OY 151 LNAVEK---IANVDEL---GPDGIDVDE--PNSGFDGLN---DKEKADFVQVYTK- 197  
 DB 171 ADPVARGNFAASVBEFLKRYGFDVLDMEYGVSGGLPGNSRPRDKRNYTLLOECVKN 230  
 OY 198 -----LREYCDKLISISQS-----SNGALSCIG----- 222  
 DB 221 IMLQKQRTAREYLLTIASGASPEYVSNTELDKIQTQVIMINIMTYDFNGQSIHANAP 290  
 OY 223 -FNPD-KKICMDDEAPYN---SKYFNKPDYKELL-----RAAQMASAG- 262  
 DB 291 LEYDPRKAEAGVPNAETNIENTVYKRYEAGYKGLVIGTPEYGRGWSGSGSGHGEYQ 350  
 OY 263 -----AYILMNNK-----DMLDMVFOVFNTNSTDSVYMKEL 296  
 DB 351 KCGPAKEGTWENGVPFSDLEKNYVNGYKRYWMDQAKVPPL--YNAENGNFIT----- 403  
 OY 297 YDSYAYYKRYDYVIMGFT--LMFPSTPEPNCKMLVYSIGDF--VTE--NKLNRRA 349  
 DB 404 YDDEGSFGHKRDFIKANLSCGMEWDFSGDSNPTLLNKLAEFKRCTETEARLNRLHH- 462  
 OY 350 DGFGLMSL 357  
 DB 463 ----LMMV 466

RESULT 10  
 O9FAC8 PRELIMINARY; PRT; 674 AA.  
 AC O9FAC8;  
 DT 01-MAR-2001 (Tremblrel, 16, Created)  
 DT 01-MAR-2001 (Tremblrel, 16, last sequence update)  
 DT 01-JUN-2001 (Tremblrel, 17, last annotation update)  
 DE CHITINASE B.  
 GN CH1B.  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1396;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-CH;  
 RA Mabuuchi N., Araki Y.;  
 RT Chitinase.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB041932; BAB16891.1;  
 DR InterPro: IPR001919; CBD\_2.  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR InterPro: IPR001412; RNA-syn\_1.  
 DR Pfam; PF00553; CBD\_2; 1.  
 DR Pfam; PF00041; fn3; 1.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR SMART; SM00060; FN3; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; UNKNOWN\_1.  
 KW Hydrolyase.  
 SQ SEQUENCE 674 AA; 74262 MW; 66BE9FE80D680561 CRC64;

Query Match 7.4%; Score 146.5; DB 2; Length 674;  
 Best Local Similarity 17.1%; Pred. No. 0.035;  
 Matches 88; Conservative 82; Mismatches 117; Indels 227; Gaps 20;

OY 6 KYSLVLCILCLCL--ISTYVSVEIGHRRAPGESRKNPREIITKFKESGKIIGYPSW 62  
 DB 5 KFTLLSLLLFLPLFLINFLINFLNALADSPKQSOX-----IYGTFPSW 48  
 OY 63 VSR--NINLKDLPNPNLVNHSFAKMDL-----SYDSIESIVGSPL 101  
 DB 49 GYGRNVOVADIDAS-KLTLNLAFAADICMNGKHGPNSTHPDNPKNQYWNKES--GYPL 105  
 OY 102 LFFSL-----IGLEYIGLNEY-----FNDAMNLRKARPDITMLLS 136  
 DB 106 QNKEVPNGTLYLGPMADYTSYSGSGTTWEDCKYACNGFGLKRLKAKPIHKTITIS 165  
 OY 137 LGGETYHPSFSDALNVEKIANLVD-----LGFQDIDVDE-----PNSGFDG 181  
 DB 166 VGGWTW--SNFSDMADEKTRKFAESTVAFLRAYGFDGVDLMEYGVETIPGSR-- 221  
 OY 182 LNDKERADFVOYVYTKLREYMKDCKLISISSNGALSCIGFNDPKKICMDEAPYNSKY 241  
 DB 222 -RPEDKONFL-----LQDVARNALNKAAGDQKQYLTLTASASQRY 263  
 OY 242 FNRPDYKELLRAQAOMASAGAIYLMNLRKMDIMVFOYTEN-----TNSDSTWKE 295  
 DB 264 ADHTELKK-----ISQILDMINIMTYDFHGGWATSNHNAALYKD 303  
 OY 296 LYDSYA-----YGKRY----- 307  
 DB 304 PNDPAANTNFHVDGAINVYTNVGVVDKLYGVFPYGRGMRKSCGKENGQYQCPKPSDG 363  
 OY 308 -----DYIIMGFTLMPST-----PNNP 326  
 DB 364 KLASKGWDDYSTGDTGVYDGLAANTVKNRGFVRYMNDTAKVYLYLNATGTFISYDD 423  
 OY 327 NDRMLVKSIGDFVTEKNLKRADGFLMSLSD 360  
 DB 424 NESMKYKT--DYIKT-----KGLNGAMFWELSGD 450

RESULT 11  
 ID 09KT32 PRELIMINARY; PRT; 1088 AA.  
 AC 09KT32;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CHITINASE, PUTATIVE.  
 GN VC1073.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE-20406833; PubMed-10952301;  
 RA Heidelberg J.F., Eilen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
 RA McDonald L., Uitterback T., Fleischmann R.D., Niernan W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 CC EMBL; AE004188; AAP94232.1; -;  
 DR TIGR; VC1073; -;  
 DR InterPro: IPR001579; Chitinase\_2.  
 DR InterPro: IPR003610; Chitin\_bind.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 2.  
 DR SMART; SM00495; Chbd3; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; 1.  
 KW Complete proteome; Glycosidase; Hydrolase.  
 SQ SEQUENCE 1088 AA; 118409 MW; 34373BA5FA4089C CRC64;

Query Match 7.4%; Score 146.5; DB 2; Length 1088;  
 Best Local Similarity 23.6%; Pred. No. 0.07;  
 Matches 79; Conservative 60; Mismatches 103; Indels 93; Gaps 21;

OY 39 NPREIITKFKESGKIIGYPSWSYHNLK-----DLNPNLVNHSFAKMDL-----KMDL 88  
 DB 383 HPRIT-----GYTSMRTKNGLPAYLAGDLPWEKLTINAFASINKSDF 429  
 OY 89 SYDSIESIVGSPLFKSLIGLE-----YIGLNEYFNDAMNLRKARPDITMLLSIGS-E 140  
 DB 430 SMQVDD--ATKMTWENVPGAEEMPSPLYOG---HFNLISFKKQYDVTLLISVGMAE 484  
 OY 141 T--YHPSFSDA-----LNAVEKTA-NLVDELGFDGIDVDE-PNSGFDGLNDEKA 188  
 DB 485 TGGFYPMTTDLASCSVNNEGIKAFNKSADVDFIRQYDGDVIDEYPSMKDSGN--PV 541  
 OY 189 DFEVOYVYTKLREYMKDCKLISISSNGALSCIGFNDPKKICMDEAPYNSKYFNRPDYK 248  
 DB 542 DF--EQSKCKGQMDNVMVMTLKR-ALPKAGEDEGRKIMLTIASP--SSAT----- 590  
 OY 249 KELLRAQAOMASAGAIYLMNLRKMDIMVFOYTEN-----YNSDSTW 293  
 DB 591 --LRGMQDFA-----MQVLDVYNIMSYDLHGTVNFEVGPQALFDGKDAELA 638  
 OY 294 K-ELYDSYAYGKKYDYIIMGFTLMPSTPFENN 327  
 DB 639 KMGVYTTAETGIGY--LNOAWTHHFFGAFKPS 670

RESULT 12  
 ID 09PAQ1 PRELIMINARY; PRT; 560 AA.  
 AC 09PAQ1;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CHITINASE.  
 GN CT51.  
 OS Ajellomyces capsulata (Histoplasma capsulatum).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Onygenales; Onygenaceae; Ajellomyces.  
 OX NCBI\_TaxID=5037;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G217B;

RA Thomas P.W., Cole G.T.;  
RT "Identification and functional characterization of a chitinase from  
RT Histoplasma capsulatum: protein purification and isolation of genomic  
RT and full-length cDNA."  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL: AF159366; AAF80370.1; -  
DR InterPro: IPR001579; Chitinase\_2.  
DR InterPro: IPR001223; Glyco\_hydro\_18.  
DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
DR PROSITE: PS01095; CHITINASE\_18; 1.  
KW Glycosidase; Hydrolase.  
SQ SEQUENCE 560 AA; 61347 MW; EF34B32ED9B064E2 CRC64;

Query Match 7.4%; Score 146; DB 3; Length 560;  
Best Local Similarity 21.7%; Pred. No. 0.029; Indels 122; Gaps 22;  
Matches 89; Conservative 56; Mismatches 143;

QY 32 RPSGRKNPREIITFKESGIIQGYPSWVS--NHNLKLPNLNVHMSFAKMDLS 89  
DB 140 RSPSPSPSHMSS-PDGYSIV--YVNMATYARNYDQL-PVKKLTHTVA----- 190  
QY 90 YDSIESIVGSLFKSLGLE-----YIGLNEY--FNDANLKRARDIIMLSLG 138  
DB 191 FANVAEAGEVELDTVADTKHPTDSWSEFGNNVGVCKQLFLKKNRHLKYLSTIG 250  
QY 139 GETYHPSFSDGALNA-----VEKIANLVDELGPDGIDVYE--PNSGFDLNLKEKADF 190  
DB 251 GWITSP-HFGAAVSTPAARTKFAFSATQNLNLGPDGIDIMEPT-----KDEEAKS 302  
QY 191 FVOYVTKREYKCDKLISISSNGALSCIGFNDPKKICMDDEAPYNSKYFNKPDYKE 250  
DB 303 LVELLKTRE-----VLDLAGKD-RREFLTVCAPGRQNEK----- 339  
QY 251 LLRAQAASAGAIYLLNNKMDIMFVOTFNTNSTDSIV----- 292  
DB 340 -LRREMT-----PYLDFYMLMAYDYSGWSDTLAGHQSNIKSNKSTPF 385  
QY 293 -MKELYSYAYGKKYDVIIMGFTLM-----FPSTPPNPN--DKMLKSIQDF--- 338  
DB 386 STRAVIDYVGVGKPSKLLGMPLIGRTPADTDGPGTPTPHGGGCGSFERGIWDYKSL 445  
QY 339 -----VTENKLNKRADGFLMS-----LSSDNAHNSQALIEYVE 375  
DB 446 PKVGAVEHIDLEKGGCG-ASMSYDASSRTMISYDNVAMVE-KTKYTIQ 493

RESULT 13  
Q9HEM6  
ID Q9HEM6 PRELIMINARY; PRT; 493 AA.  
AC Q9HEM6;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE CHITINASE.  
GN CTS2.  
OS Ajiellomyces capsulata (Histoplasma capsulatum).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Onygenales; Onygenaceae; Ajiellomyces.  
OX NCBI\_Taxid=5037;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-G217B;  
RA Specht C.A., Benfield B.B., Garcia J.J.;  
RT "Identification of bacteria-like chitinases in fungi."  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF315588; AAG41982.1; -  
DR InterPro: IPR001223; Glyco\_hydro\_18.  
DR InterPro: IPR000531; TonB\_box.  
DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
DR PROSITE: PS00430; TonB\_DEPENDENT\_REC\_1; UNKNOWN\_1.

SQ SEQUENCE 493 AA; 53623 MW; E9C420F2790D9179 CRC64;

Query Match 7.2%; Score 142; DB 3; Length 493;  
Best Local Similarity 22.2%; Pred. No. 0.046;  
Matches 70; Conservative 43; Mismatches 102; Indels 100; Gaps 15;

QY 58 YPSWVS--NHNLKLPNLNVHMSFAKMDLSYDSIESIVGSLFKSLI----- 107  
DB 101 FYVMATYGRNYQPDQL-PAKLTHTVA-----FANVHGDTGEVYLLINDYADTDKHYPT 154  
QY 108 -GLEYGILNEY--FNDANLKRARDIIMLSIGETVHPSFSDGALNAVERKIALVYD-- 162  
DB 155 DSWEDVGNVYGVCKQLFLKKNRNLKILLISIGGWY-SSNFGAASGANRAHFAFTA 213  
QY 163 -----ELGFDGIDVYE--PNSGFDLNLKEKADFVQYVTKREYKCDKLISISSNG 216  
DB 214 TKMLMDGFDGLDIMEPT-----NDEAKNVELLKTRE-----KDELSKN--- 258  
QY 217 ALSGIFNDPKKICMDDEAPYNSKYFNKPDYKELLRAQAASAGAIYLLNNKMDIM 276  
DB 259 -----RKFTLVACAPGRKFK-----LRLEMT-----PYLDF 288  
QY 277 VFQTFNTNSTDSIV-----MKELYSYAYGKKYDVIIMGFTL 317  
DB 289 YNLMAVDYAGSMDVYAGHANLEVSKSDPKSPYSTALDYIGVGEVPAKMTLGMPL 348  
QY 318 M-----FPSTPF 324  
DB 349 YGREFADTDGPGTPT 363

RESULT 14  
P94339  
ID P94339 PRELIMINARY; PRT; 604 AA.  
AC P94339;  
DT 01-MAY-1997 (TREMblrel. 03, Created)  
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE CHITINASE (EC 3.2.1.14) (CHITODEXTRINASE) (1,4-BETA-POLY-N-  
DE ACETYL-D-GLUCOSAMINIDASE) (POLY-BETA-GLUCOSAMINIDASE).  
GN CHT.  
OS Bacillus licheniformis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_Taxid=1402;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TP;  
RA Taitlavanich S., Panbangred W.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N-  
CC ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL: U71214; AAB47847.1; -  
DR HSSP: P07254; ICTN.  
DR InterPro: IPR001579; Chitinase\_2.  
DR InterPro: IPR003610; Chitin\_bind3.  
DR InterPro: IPR003962; Phitt\_repeat.  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR001223; Glyco\_hydro\_18.  
DR Pfam: PF00041; fn3; 1.  
DR Pfam: PF00704; Glyco\_hydro\_18; 2.  
DR PRINTS: PR00014; FMYTPHIII.  
DR SMART: SM00495; ChtBD3; 1.  
DR SMART: SM00060; FN3; 1.  
DR PROSITE: PS01095; CHITINASE\_18; 1.  
KW Glycosidase; Hydrolase; Repeat.  
SQ SEQUENCE 604 AA; 66837 MW; 65B39258444BB3D2 CRC64;

Query Match 7.2%; Score 142; DB 2; Length 604;



---

**THIS PAGE BLANK (USPTO)**





FEATURES	77555-06091, USA
SOURCE	Location/Qualifiers
	1..1137
	/organism="Plasmodium falciparum"
	/isolate="K1"
	/db_xref="taxon:5833"
	/chromosome="12"
	/country="Thailand"
mRNA	<1..>1137
	/product="chitinase"
CDS	1..1137
	/codon_start=1
	/product="chitinase"
	/protein_id="AAFI6902.1"
	/db_xref="GI:5561404"
	/translation="MNFVRYSLFVLCILLCCLSTYSVSYIEGRARPGESKRKNREIT KTFESGKGIIGYYPSWVSYNHNLKDLNPLNVVHMSFAAMDLSYDSIESIVGSPILL FKSLGLEIYGLINEFYNDAMNLRKRPDIIMLSLGETYHPSFSDALNAVEKIANI VDELGFQDIDVDEPENGSGDLNDRKADPFVOYVYTKLREYMCDDKLISISONGAL SCIGNEDNRKICMDDEAPYNSKYFNKPPVKKKEILRAOMASAGGAIYLMNNLKIMDM VFOFENTNSTDSYMKELIDSYAYKQKRYDVIYGLFPMSPSTFPFNPRDKMLVNS IGDVPKTNLKNKRADGRLGSLSSDNNAAHNEQLAIEYFVESH"
BASE COUNT	321 a 263 c 257 g 296 t
ORIGIN	

Query Match	Similarity	100.0%	Score 1137	DB 3	Length 1137
Best Local	Similarity	100.0%	Pred. No. 7,1e-301		
Matches 1137	Conservative	0	Mismatches 01	Indels 0	Gaps 0
QY	1	atgaattttaccgtaaaataactcttccctggtatcttgctgctgctgctgct	60		
Db	1	ATGAATTTTACCGTAAATAACTCTTCCGTTATTGTCCTGTGCTGCTGCT	60		
QY	61	acttattgtttcgttatcgaaagtcattcgctgctccggtggaatcccgaaaaacccg	120		
Db	61	ACTTATGTTTCGTTATCGAAGGTATCGTCTGCTCCGGGTGAATCCCGAAAAACCG	120		
QY	121	cgtgaatatatcaaaaacgctttaaagaatccgtaaaagtgatcatccaggttactatccg	180		
Db	121	CGTGAATATATCAAAACGTTTAAAGAATCCGTTAAGGTATCATCCAGGTTACTATCCG	180		
QY	181	tccctggttgagactataatcacaaacctgaagaacctgaaccggaactggaacgtgttcaac	240		
Db	181	TCCCTGGGTGAGCTAATAATCACAACTCGAAGACCTGAACCCGAAACCTGAAGCTGAGTTCAC	240		
QY	241	atgaaccttgcctaaatgatttaagctacagacatccatctgataatcatgtagtagagcccg	300		
Db	241	ATGACCTTTCGCAAAATGGAATTTAAGCTACAGACTCATTGAATCCATGCTGATGATGCCCG	300		
QY	301	ctgctgctttaaaagccgtgattgacctgagatatalgtgctlgaatgagttactcaatgat	360		
Db	301	CTGCTGTTTAAAGCCCTATGTGGCCTGGAAGTATTTGGTCTGAATGAAGACTACTTCAATGAT	360		
QY	361	gccatgaatcttcgacgaagaagctcgtcccggaacattatagctgcgtgtagagctgggtgag	420		
Db	361	GCCATGATCTTCGCCGAAGCTCGTCCGCACATTAATTAACTGCTGATGAGCTGGCGGTGAG	420		
QY	421	aactacacatccgctccctctcgattctgcatctctgacactcaacggtgtgaaaaaatcgcgaaltcg	480		
Db	421	ACCTACCATCCGCTCCTCTCGATTCTGCATCTCACTCAACGGGTGAAAAAATGCAAAATCTG	480		
QY	481	gtgagatgacacgggcttcgatggtatgtatgatatcagaacccaagaggaagctttgat	540		
Db	481	GTTGATGACACGGGCTTTCGATGTGATTATGATGATGATTCGAAACCGAAGGGAGCTTTGAT	540		
QY	541	ggtctgaaatgataaagaanaaagcgacctctttgtacataatcgtyaaccaactcgcgaa	600		
Db	541	GGTCTGAATGATTAAGAANAAGCGACCTCTTTGTACATAATCGTACCAAACTCGCGCA	600		
QY	601	tacatggtgatgataaactgatcaactgaactcgaagctctcttaatgycgtctgagctgac	660		
Db	601	TACATGTTGATGATTAACCTGATCACTGATCTGACCGAGTCTCTTAATGCGGCTGTAGCTGC	660		

QY	661	atcgggtttaacgaccccggaaaaaactcttatgagtgaagaagcttcggtataaacgcaa	720
Db	661	ATCGGTTTAAAGCACCCGGAATAAAATCTATATGAGTGAAGAAAGCTCCGTATAACGCNAA	720
QY	721	tattcaacaaacccggaagcttaagaaagactgttaacgcgaccccgataagcttcg	780
Db	721	TATTTCAACAAACCCGAGCTTTAAGAAAGAACTGTATACGGCGACGCCAGATGGCATCTGGG	780
QY	781	ggtgtgtgccatctaccctgtatgaacaacctgaagaatataatgtcaatgtgtgtgcag	840
Db	781	GGTGTGGCCATCTACTCCTGTATGAACAACCTCGAAAGATATGATTGACATGGTGTGGTGCAG	840
QY	841	agcttacaataacgagacacttaacgaatttaccggtatgaagaatataagaaacctac	900
Db	841	AGGTTCAACTACACGACACTCTACCGATTTTACCGTTATTAAGAAATTAATACACTCTAC	900
QY	901	gcatactatgttaaaagagctaatcgcgtatcatcatatgtgccttcaacctgtatgttc	960
Db	901	GCATACTATGTATAAAAGTAAACGATTTACGTATCATCATCATGGGCTTCACCTGATGTTCCG	960
QY	961	tccaagccgttcaaccccggaacgataaataatgcctgtgttaaaatctatcgtgcgatttcg	1020
Db	961	TCCACGCGCTTTAACCCCGAACGATTAATAATGCTGTAAATCTATCGCGCATTTTCGTAAAA	1020
QY	1021	acttgaacaaacactgaataacgcgcagatgtgttcctgcgcctgtgtctgtgccagcgac	1080
Db	1021	ACTGAAACAAACTGAAATTAACGCGCGAGATGGTTTCGGCTGTGGTCTGTGCCAGCGAC	1080
QY	1081	aacgcgcgccataatgaacagctgtgcgactcgaagctacttcgtttgaaaagcctgtca	1137
Db	1081	AACGGCGGCCAATATGAACAGCTGGGAGTGGAGTACTCTGTTGTAAGAACCTGTGCATTA	1137

	RESULT	2
AC006279/c	LOCUS	
DEFINITION	AC006279	245802 bp DNA linear HTG 12-AUG-2000
ACCESSION	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces.	
VERSION	AC006279	
KEYWORDS	AC006279.8 GI:9797723	
SOURCE	HTG; HTGS_PHASE1.	
ORGANISM	malaria parasite P. falciparum.	
REFERENCE	Plasmodium falciparum	
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
TITLE	1 (bases 1 to 245802)	
JOURNAL	Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W.	
REFERENCE	Plasmodium falciparum 3D7 chromosome 12 unpublished	
AUTHORS	2 (bases 1 to 245802)	
TITLE	Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B. and Davis,R.W.	
JOURNAL	Direct Submission	
AUTHORS	Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	
COMMENT	On Aug 12, 2000 this sequence version replaced gi:8810451.	
	* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	
	* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
	* 1 147568: contig of 147568 bp in length	
*	147569 147768: gap of unknown length	
*	147769 147769: contig of 69702 bp in length	
*	217471 217670: gap of unknown length	
*	217671 219372: contig of 1702 bp in length	
*	219373 219572: gap of unknown length	
*	219573 221333: contig of 1761 bp in length	
*	221334 221533: gap of unknown length	

[illegible]

QY 154 aaagratatccaagggtactactccgtcgtgggtgagctataatcacaccgtgaagac 213  
AF216703 1 AAAGGAATTTATTCAGGTATATATCATCATGAGTTTCTTATTAATCATTAACCTTGAAAGAT 60  
QY 214 ctgaaccggaacctgaacgtggttcacatgagcttgcctaaatgattgaatgaacac 273  
Db 61 TTAATATCTTAATTAATGTTGTACATATGCTTTTGCTAAGATGATTTATCTATAC 120  
QY 274 tccatgaatccatcgtgagtagcccgctcgtctgtttaaagccgtatggccctgagatc 333  
Db 121 AGTATAGAGATATGTTGGTCTCCCTTACTTTTCAAGTCATTAATAGGTTTGAAAT 180  
QY 334 attgctgtagatgaacttcaatgatgcctatgacgtgcgaagccgtccgagacat 393  
Db 181 ATAGGGTTAAACGATATTTTAATGATGCCATGATTTAAGAAAAGCTCCGACGATATT 240  
QY 394 attatgctgtagacgtcgtgagcggtgagaacctacatccgtccctccttcgactc 453  
Db 241 ATTATGCTTTTATCTGTGAGGAGAAACATATCATCAAGTTCATTTGATTCGCTTTA 300  
QY 454 aacgcggttgaataaataacgcgaatcgtgtgatacgtgagctcgtatgattgata 513  
Db 301 AATGCTGTTAGAAAATTTGCAAAATTTAGTTGATGAATTTAGATTTGATGAAATGAT 360  
QY 514 gattacgaaccggaacgagcttgcgtgctgatagtataaagaataaagagactctt 573  
Db 361 GATTATGAGGCTTAATGATCTTTGATGCTTTGATGAAACGATTAAGAAAAGCTGATTTTTC 420  
QY 574 gtacaatacgtlgaacaaactlgaacatacgtlgtgatagtataaactlgaacagc 633  
Db 421 GTCAATATGTTAACAATAATTTGAGGAAATATATGCTGATGATTAATTAATTTCTATATCT 480  
QY 634 cactccctatagcgtcgtcgtgagctgagctcgtgttcaacgacccgaaataatcgtatg 693  
Db 481 CAATCATCTAAGGAGGCTTTAAGTTGTATAGTTTATGATTCCTAAGAAAATATATATG 540  
QY 694 gatacgaacgtccgtlgaacgaacaaatattcaacaacacggaacgtlgaagaagaacgt 753  
Db 541 GATGATGAACGACCATTAATTAATCAAAATTTTCAATAAACCCTGATGTAAGAAAGAGCTA 600  
QY 754 ttacgcgacgccaagatgagctcgtcgtgaggtgagtcacatcgtatgatacaacctgaa 813  
Db 601 TTAAGACGACCGCAAAATGGCTTCAGCTGAGAGACCAATATATTTAATGAATTAATTAAG 660  
QY 814 gatacgtlgaacgtggtgtlgtgcagagctlgaacgaacgaacgaacgaacgaacgaac 873  
Db 661 GATATGATTTATATGTTGTTGTACAAACCTTTAATTAATTAATTAATTAATTAATTAATTA 720  
QY 874 gttatgaagaatatacgaactcctacgaactatagtataaagaatgaactgaactga 933  
Db 721 GTATGAAGAATAATTAATGATTTCTTATGCTTATATGAAAAAATATGATTAATGTTT 780  
QY 934 atcaatggtcctacacgtatgctcgtlgaacgacgctgaacacggaacgaactgaactg 993  
Db 781 ATATGATGATTTATCATTAATGTTCTTCAACCCCTTTTAAATCGAATGATTAATGCTT 840  
QY 994 gtaaaatcctcgtgaggttgcgtataaactgaacgaacgaacgaacgaacgaacgaacgt 1053  
Db 841 GTAAATCCATTTGGTGTATTTTGTAAAGACTGAATAATTAATTAATTAATTAATTAATTAATTA 900  
QY 1054 ttgcgcctggtcgtcgtcgtcgaacgacgaacgacgaacgaacgaacgaacgaacgaacgt 1113  
Db 901 TTTGATGATGATGCTTATATCTTCTGATTAATGCTGCACATATTAATGAATTAATGAATTA 960  
QY 1114 tactcgtlgaagaacctgcactaa 1137  
Db 961 TATTTCTGTAAGATCTTACATTA 984

RESULT 4  
AF216703 1764 bp DNA linear SYN 26-JUN-2000  
LOCUS AF216703

DEFINITION Synthetic construct chitinase CHT1 (Chl1) gene, complete cds.  
ACCESSION AF216703  
VERSION AF216703.1 GI:8705223  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
SUBMISSION  
DIRECT SUBMISSION  
14-DEC-1999 WHO Center for Tropical Disease, University  
of Texas Medical Branch, 301 University Blvd., Galveston, TX  
77555-0609, USA

FEATURES  
source  
1. 1764  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Plasmodium gallinaceum gene engineered for  
bacterial codon usage"  
1. 1764  
/gene="Chl1"  
1. 1764  
/gene="Chl1"  
/codon\_start=1  
/transl\_table=1  
/product="Chitinase CHT1"  
/protein\_id="AAE78777.1"  
/db\_xref="GI:8705224"  
/translation="MNFKISIFLIIVSILYSANSRTLRKNNINSLGIRNNKTH  
OTFIHSEFSLKSNNSNFEYEGYCGDGNSTLRNNKNNINNDKSPROILEYKR  
KOGITAGYCSNMSGDRKHMDSNPTSLIYAFARNIMLYDSRPNGBRELLR  
KGLIETTYGMMLETRIRKRPVPIIILSLGETYMDIKEYDIDYDKILKLVND  
DLGVDIDWEPHKEFYNLELNSNYIKLINDLRTPEERLISISGSSNAALCSVS  
GVASFCDKDESPYNTKFLSEQATETNNELRAAMLSAGFINIFMTAEKIDLVLIQT  
YLNLETNPDIWDMVLYSHLYFLKYNITIIIGLSLENNRNGSPSPENKLELVTIOT  
DKNONNRADGIGIMHLEPMKEOLPTGSPVDIFLTIYMKHNPVOTPDULITENPE  
DCSTIDEVYGVIVPIIGIYKKNDAIMKTRSYSHAPVDREMDLVKVCYKICDG  
KAHHYNTDKRESITIKGPEPLIMMOOGPEEGALESTYKLDASKCPGIEBNKK  
YPRKPLEVEBYEVDLPQ"

BASE COUNT 545 a 422 c 386 g 411 t  
ORIGIN

Query Match 15.18; Score 171.6; DB 12; Length 1764;  
Best Local Similarity 52.28; Pred. No. 5.6e-36;  
Matches 512; Conservative 0; Mismatches 444; Indels 24; Gaps 5;

QY 108 ccgcaaaaacccggtgaattatcaaaacgtttaagaatccgtaagaagtatccca 167  
Db 264 CCGCAAAAGCCCGGCCAGATTTCTGAAGAATACAAAACGCAACAGGCGATTATTGC 323  
QY 168 ggttactatccgtcctggtgagctataatcaca-----accatgaaggaacctgaa 218  
Db 324 CGGCTACTATGTTCTCGGAACACGACGAGGCGCAACCGTCAACACATGATGCATCTAA 383  
QY 219 ccggaacctgaagtggtlcaatgaagcttgcgtataaagtattgaactgaagctccat 278  
Db 384 TCCGATGATGTCACCTCTGATATCGGCTGCCCGGATTAATATCTGATATGACCTTTC 443  
QY 279 tgaatccatcgttagtagtgcgcgcgtgtltaagaacctgaactgtgagctgagatattg 338  
Db 444 TCGTCCGTTTAACGCTGCGACGCTTTTCTGCTGCGTAAACATAGGCTGGAATACGAAC 503  
QY 339 tctgaatgagttactcaatgatgacatgaatcgtgcgaacacgtcgtccgagacattat 398  
Db 504 CTATGCGATGATGCTGATGAATTAATTCGTCGATCCGAAAGTACGCGGACGATTAATCA 563  
QY 399 gctgcagacctgagcggtgagacattacatccgtccctccttcgattctgacactcaagc 458

Db 564 TCTGTTAACTTAGGGGGGAAACCT--ATATGATCGACATCGAAAAAGAAATCGACTA 620  
 Qy 459 ggttgaataaatccgaactctggtgataactgagggcttcgatagtatgata 518  
 Db 621 TGTGATTAATAATCTGAACACTGTGTGAACGATTGTGATCGATGACATCGACTG 680  
 Qy 519 cgaacggaacgagctctgatagtctgataaagaagaaggagctcttgata 578  
 Db 681 GGAAACCGCATGGCAAGTTTACATCTGATGATGAACTGAACTTCAGCAATCTATATCAA 740  
 Qy 579 atacgagcaaaactcgcgaataacatgctgataaactgatacagatcagacgcagc 638  
 Db 741 ACTGATTAACTCTCTCGGAAACCAATCCCGGGAAGAAACTGATCTCTCTGCTG 800  
 Qy 639 ctctaaatgagctctgatactgatacgtttcaacagacccggaataaatctgata 698  
 Db 801 CAGCAACGAGACCTGAGTGGGTGAGCGGCGCTTC-----TTCTCGCAAGATGA 854  
 Qy 699 cgaagctccgataaagaacaaatattcaacaacc---ggaagcttaagaagaactgtt 755  
 Db 855 AGAATCCCGTATTAACACTAAATTCCTGAGCGAACAATCGAAACTAACAAAGAGCTGCA 914  
 Qy 756 acgagccacacagatgacatcgcgagtggtgagctacacacgataaactgaaga 815  
 Db 915 TGTGCTGCGCGCAATGCTGACGCGCGTACGTTCAATTAACATTTTCAACACTCGCAAGA 974  
 Qy 816 tatgatacagatggttctgacagactcactacacgaactcactacatctacgt 875  
 Db 975 AAAAATCGACGTGATTTATCCAAAC--CTATATCTGAGAGACCAACCGCATAT 1031  
 Qy 876 tatgaagaatatacagactcactacgataactatgataaagaactacgatacgt 935  
 Db 1032 TATGTCGATGATGATCTGAGCGACCTGTATTCGGCTCGAAATACAACTTACGATCAT 1091  
 Qy 936 cactgagctcaccacgagatgtccgcgcacgcgcgttcaacccgaagaataaagtgtgt 995  
 Db 1092 CCTGGTTTTCTCTGGAACAACAATCGTGGCGTTTACCGCGGAACAAGAACTGCT 1151  
 Qy 996 aaatcatalcgcgacttctgataaactgaacaaactgaataaagcgcagatggtt 1055  
 Db 1152 GGAATGTGTTGGCAAAACATATCCAGCAAGAAACCGAATATTAACCGGTGATGCTAT 1211  
 Qy 1056 cgcgcgtggtctctgtcca 1075  
 Db 1212 CGGTATTTGGCATCTGTCTCA 1231  
 RESULT 5  
 AF064079 2500 bp mRNA linear INV 10-APR-2000  
 LOCUS AF064079 Plasmodium gallinaceum endochitinase precursor, mRNA, complete cds.  
 DEFINITION AF064079 Plasmodium gallinaceum endochitinase precursor, mRNA, complete cds.  
 ACCESSION AF064079  
 VERSION AF064079.1 GI:7530423  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Plasmodium gallinaceum.  
 Plasmodium gallinaceum  
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 REFERENCE  
 AUTHORS  
 Vineitz,J.M., Valenzuela,J.G., Specht,C.A., Aravind,L., Langer,R.C.,  
 Ribeiro,J.M. and Kaslow,D.C.  
 Chitinases of the avian malaria parasite Plasmodium gallinaceum, a  
 class of enzymes necessary for parasite invasion of the mosquito  
 midgut  
 JOURNAL J. Biol. Chem. 275 (14), 10331-10341 (2000)  
 MEDLINE 20209408  
 REFERENCE 2 (bases 1 to 2500)  
 AUTHORS Vineitz,J.M., Valenzuela,J. and Kaslow,D.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAY-1998) Pathology, University of Texas Medical  
 Branch at Galveston, Kellier 2.138, 301 University Blvd, Galveston,  
 TX 77555-0609, USA  
 FEATURES  
 source 1..2500 Location/Qualifiers

/organism="Plasmodium gallinaceum"  
 /db\_xref="taxon:5849"  
 /dev\_stage="ookinete"  
 276..2039  
 /EC\_number="3.2.1.14"  
 /codon\_start=1  
 /product="endochitinase precursor"  
 /protein\_id="AA063208.1"  
 /db\_xref="GI:7530424"  
 /translation="MNFKISFLILVSLIYSANSPRLGKNNINNSLITRENNKTH  
 OTEHESFSLKSNNSNFEVSGDGCNSRITKNNNNINKNDKSPDILEYKKR  
 KOGITAGTYGSGNSGDRAKHIDNPVSIITYAPRIIMLDVDSRPNGQRLR  
 KHGLEIYEGMMLNEIRIRKRPVYIILSLGSEYIMIDEKEIDYDKILKLVND  
 DLGDVIDMEPRGKYNELNEFSNYIKLINLKRITPEKRLISIGSSNAISVS  
 GVASFCDEESPYNTKFLSEOTKEKHELAAMLSACTIINIFMTAKELIDVFIQT  
 YNLETTNDIMVDYLSHLRGLKNTYIILGSEFHNHNGGFSPEKLELIVGTIH  
 DKNONNNRADIGITWHLRMDLPDPSGDVIDFELNINKLNHNPQEPKDLITFENP  
 DCSTIDEVPRGLVIPTIGIYYKHNDIAIKTRSYISHAAGVDREMDLYKVCYKICDG  
 KAAHYNTDYKSSIIIMKGEYLIKMMQGGPEEQALESTYTKLDASKCPGLEENKK  
 YPKHLEVEEDYEQVDPLO"  
 BASE COUNT 993 a 283 c 365 g 858 t 1 others  
 ORIGIN  
 Query Match 9.2%; Score 104.2; DB 3; Length 2500;  
 Best Local Similarity 48.0%; Pred. No. 1,8e-17;  
 Matches 468; Conservative 0; Mismatches 483; Indels 24; Gaps 5;  
 Qy 110 gcaaaaaccggtgaatatacaaacgtttaagaatccggtlaagatataccag 169  
 Db 541 GAAATACACCAAGACAAATTTTGAAGAGATATATAAAGAAACAAAGATTTATGACG 600  
 Qy 170 gttactatcgcctcgtggtgagctataatcaac-----acctgaagacctgaacc 220  
 Db 601 GATFACATGTTCTATGACGACGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 Qy 221 cgaacctgaagctggttcaatgagcttgcataaattgaattgaactgaactcatgt 280  
 Db 661 CATGGTGTCAATTTATATATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Qy 281 aatccatcgttaagtgccgcgcgtttaaagaacctgttgccggtgatatgtgc 340  
 Db 721 GACCATTTAATGAGACAAAGATTCCTATTAAGAAACACGCGCTTAAGATGAACCT 780  
 Qy 341 tgaatgagctactcaatgcatgataatcgcgaacgtcgcgcgacatataatgc 400  
 Db 781 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
 Qy 401 tgcagagcctggtggtgagacctataccatcgctccctcgattctgactaaccgag 460  
 Db 841 TTTTATCTTATGAGTGGAGAACT--ATATGATGATGATGATGATGATGATGATGATG 897  
 Qy 461 ttgaaaaaatcgaactcgtggtgagacctgaactgagctgatagtatgataagtaag 520  
 Db 898 TGGATTAATAATTTGAAGCTGTGATTAATGATTTGATTAAGTGTAGATTAAGTGGG 957  
 Qy 521 aacggaacgagcttgaatgctgataaagaagaagaacgagctcttgaacat 580  
 Db 958 AACCAACATGGAAGATTTTCAACTTAATGAATTAATTTTCAATTAATTTTCAATTAAT 1017  
 Qy 581 acgtgaccaactcgcgagatacatgctgataaactgaactgaactgaactgaactcct 640  
 Db 1018 TAATTAACCTTTTAAAGAAACCAATTCCTGAGAAAGTAAATTTTCAATTTTGTGTCAT 1077  
 Qy 641 ctaatgagcctcgtgagctgacatcgcttcaagaacccggaataaatctgataagtaag 700  
 Db 1078 CAATGCTGCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1131  
 Qy 701 aagctcgtataaagcaataattt---caacaacacgagcgtlaagaagaactgttac 757  
 Db 1132 AATTCGACATATTAACCTAATTTTGTGTGACCAATATTAACCAATTAAGATTAACATA 1191  
 Qy 758 gcgcagcccaagatgacatcgcggtggtgcatctactcgtatgaacaacactgaagaata 817

Db 1192 GGGCAGCAGCATGTATACACAGAACTTTATTAATTTTATTAATACAGCAAGAGAGA 1251  
QY 818 tgatgacatggtgttctgtgagacggttcaactcaacgaactctacgattctaccgtta 877  
Db 1252 AAAATAGATCTTGTATTTATTTAAACATACAAATTT--TAGAAACATACAAATCCAGATATTA 1308  
QY 878 tgaagaattatacgaactcctacgatactatggttaaaagtaacgattacgattgacatca 937  
Db 1309 TGGTAGATATCTACTTATCCATTTATTTGTTTAAATATATACATCACAATCATAT 1368  
QY 938 tgggtctacccctgatttcccgctccacgcttcaacccgaacgaataaagtctgtaa 997  
Db 1369 TAGGTTTTCATATGAAACATACAGAGGTGATTTAGTCCGAAATATTAAGATATTTAG 1428  
QY 998 aatctacgagatttctgtaaaactgaaacaaactgaataacgacgagatggttcg 1057  
Db 1429 AATTGATAGGAAACATACATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 1488  
QY 1058 gacctggtctctgt 1072  
Db 1489 GGAATATGCAATTTAT 1503

RESULT 6  
PBE305256 2562 bp mRNA linear INV 30-MAY-2001  
LOCUS Plasmodium berghei mRNA for chitinase (chl gene).  
DEFINITION AJ305256  
ACCESSION AJ305256 GI:14275848  
VERSION  
KEYWORDS chitinase; Chl1 gene.  
SOURCE Plasmodium berghei.  
ORGANISM Plasmodium berghei  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
REFERENCE 1 (bases 1 to 2562)  
AUTHORS Dessens,J.T., Mendoza,J., Claudianos,C., Vinetz,J.M., Khater,E.,  
Hassard,S., Ranawaka,G.R. and Sindén,R.E.  
TITLE Knockout of the rodent malaria parasite chitinase pbchl1 reduces  
infectivity to mosquitoes  
JOURNAL Infect. Immun. 69 (6), 4041-4047 (2001)  
2 (bases 1 to 2562)  
Dessens,J.T.  
REFERENCE Direct Submission  
AUTHORS Submitted (31-JAN-2001) Dessens J.T., Dept. of Biology, Imperial  
College of Science, Technology and Medicine, SAFB, Imperial College  
Road, London, SW7 2AZ, UNITED KINGDOM  
FEATURES  
source location/Qualifiers  
1..2562  
/organism="Plasmodium berghei"  
/strain="ANKA"  
/isolate="2.34"  
/db\_xref="taxon:5821"  
174..227  
/gene="chl1"  
174..2120  
/gene="chl1"  
/function="mosquito infection"  
/codon\_start=1  
/product="chitinase"  
/protein\_id="CAC40151.1"  
/db\_xref="GI:14275849"  
/translation="NNIYVSLFLIFSTINNAYSVLGKGTITNHPINISHTEKOYKS  
LSHYDALGDSKRYVCEETKYICKNGSGEYCSFGSGNGNGVNGNGVNGNGVNGNG  
GAENGNGAENGNGVNGNGAENGNGAENGNGAENGNGVNGNGNGNHERKSPREIL  
EYKRRKQGIAGYGSWNGSGAGKQMTSHNPVSLIYIAFAIINNAYASKRYNEL  
OKFLQKRGHLEFETGMAMANEIKIRKARPDVLIISLGETVMIIDITKEIDYIPQIV  
NNVNNEDLDGVDIDEMPNKRYHLNELAFSNYIKLILRSNLSKDLNLSIGSSNA  
ALSCVSWETFECDDEFPYNTOLSEOMIVNPELYKATSTMSTGTFNVNFTAKDKID  
LYFIQYLNLETSPDIDIMDYLSHLYFGKRYDITVPLGEISIEONRGFSPDSKRYVL  
VSKAIDNNHFEHRADGVGIMHLPKQEMNGAIDLDGFIIDNTWTHLNFPIKSTLDIV  
SLPNPDSCSTIEEYVGVVPTVGTAKHNDALWVRSYSTRAFGIDRYEMNVLKQVY  
EKACNNEAAHYNVDPYPTGKIIWKGEVFSIKQMHGPGEPYLESYEKLEDEYSCPL

AEWNKHPKPIEVAPEQELD"  
gene 174..2120  
/gene="chl1"  
mat\_peptide 228..2117  
/gene="chl1"  
/product="chitinase"  
BASE COUNT 1001 a 323 c 416 g 822 t  
ORIGIN

Query Match 6.28; Score 70.2; DB 3; Length 2562;  
Best Local Similarity 45.48; Pred. No. 3.8e-08;  
Matches 467; Conservative 0; Mismatches 538; Indels 24; Gaps 5;

QY 110 gcaaaacccgcygaaatatacaaaagtttaagaatccggttaagatcatccag 169  
Db 631 GAAATCTCCAAAGGAAATTTTAGAGAAATTAATAAAGAAAGCAAGATATATTGACG 650  
QY 170 gttactcgttccctggtggaactataacacacccgaaagacgt-----aac 220  
Db 691 GATATTATGATCATGGAATAGCAAGAGAGCTAGCGGAAACAAATGATCATTCACACC 750  
QY 221 cgaacctgaacgtggttaccatgaagcttgcataaagatgaattaaagtaactccat 280  
Db 751 CACAGTATCTATATATTTATATACCTTCGCCCTATTAATATGAACTATGATGCTTCTA 810  
QY 281 aatccatgtagtagcccgctgctgtttaaaagccgtagtgccctgagatataatgctc 340  
Db 811 AACCATATATGCACTTCAAAATTTTACAAAAGCAATGACTGATTTGAAACAT 870  
QY 341 tgaatgacttcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 400  
Db 871 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930  
QY 401 tgcgaagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 460  
Db 931 TTTATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 987  
QY 461 ttgaaac 520  
Db 988 TTCTCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047  
QY 521 aacgaaacggaacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 580  
Db 1048 AACCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107  
QY 581 acgtgaaac 640  
Db 1108 TGATTTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167  
QY 641 ctaatgagcgtcgtgagctgacatcggttcaacgacccgaaacacacacacacacacac 700  
Db 1168 CAATTCACGATTTATCAT-----GTGATACGATGATGATGATGATGATGATGATGAT 1221  
QY 701 aagctcgtataacagaac 757  
Db 1222 AAAGCCATATATATACACAGATTTATTCGACAAATGATGATGATGATGATGATGATGAT 1281  
QY 757 ggcgaagccgaatgagctggtggtggtggtggtggtggtggtggtggtggtggtggtg 817  
Db 1282 AAGCATCAACCATTTTATCAACAGAACATTTGTTATGTTTATTAATCAACAAAGGATA 1341  
QY 818 tgatgacatggtgttctgtgagacggttcaactcaacgaactctacgattctaccgtta 877  
Db 1342 AAAATGATCTTGTATTTATTTCAACATATATATTTAGAAACATCAAAACCGAGAT---TCA 1398  
QY 878 tgaagaattatacgaactcctacgatactatggttaaaagtaacgattacgattgacatca 937  
Db 1399 TGGTAGATATCTACTTATCCATTTATTTGTTTAAATATATACATCACAATCATAT 1458  
QY 938 tgggtctacccctgatttcccgctccacgcttcaacccgaacgaataaagtctgtaa 997  
Db 1459 TAGGTTTTCATATGAAACATACAGAGGTGATTTAGTCCGAAATATTAAGATATTTAG 1518









```
misc_feature      /clone.lib="RPCT-11 Human Male BAC"  
                  1..1020  
misc_feature      /note="assembly_fragment"  
                  1121..2314  
misc_feature      /note="assembly_fragment"  
                  2415..2871  
misc_feature      /note="assembly_fragment  
                  clone_end:r7  
                  vector_side:right"  
misc_feature      2972..5093  
misc_feature      /note="assembly_fragment"  
                  5194..6740  
misc_feature      /note="assembly_fragment"  
                  6841..10535  
misc_feature      /note="assembly_fragment"  
                  10636..14161  
misc_feature      /note="assembly_fragment"  
                  14262..18606  
misc_feature      /note="assembly_fragment"  
                  18707..25158  
misc_feature      /note="assembly_fragment"  
                  25259..31210  
misc_feature      /note="assembly_fragment"  
                  31311..37297  
misc_feature      /note="assembly_fragment"  
                  37398..44148  
misc_feature      /note="assembly_fragment"  
                  44249..49736  
misc_feature      /note="assembly_fragment"  
                  49837..59810  
misc_feature      /note="assembly_fragment"  
                  59911..68630  
misc_feature      /note="assembly_fragment"  
                  68731..78807  
misc_feature      /note="assembly_fragment"  
                  78908..88451  
misc_feature      /note="assembly_fragment"  
                  88552..98431  
misc_feature      /note="assembly_fragment"  
                  98532..111619  
misc_feature      /note="assembly_fragment  
                  clone_end:SP6  
                  vector_side:right"  
misc_feature      111720..126345  
misc_feature      /note="assembly_fragment"  
                  126446..158515  
misc_feature      /note="assembly_fragment"  
                  /note="assembly_fragment"
```

Query Match	3.5%	Score 40.2	DB 2	Length 158515
Best local Similarity	56.4%	Pred. No. 8.9		
Matches	75	Conservative	0	Mismatches 58
				Indels 0
				Gaps 0
QY	14	taaaatactcttcctcgtgtaattcgtctcgtcgtcgtctactatggttcg	73	
Db	147413	TACATTTGGTTCTCTTTTTCAGTTCACCTGGCTTCCACCCCTTGTATTGTTTCG	147472	
QY	74	ttatcgaaggatcatcgtctcgtccgggtgaaatcccgcaaaacccgctgaaatata	133	
Db	147473	CTTCTTTTTCACCCCTTTCATCGGAGAACTTTCACATAAATTCACATAGAACGATGG	147532	
QY	134	aaacgtttaaga	146	
Db	147533	AAATATTTTAATA	147545	

RESULT 12  
AC084884/c  
LOCUS AC084884  
DEFINITION Oryza sativa chromosome 10 clone OSUNBA0031A07, complete sequence.  
ACCESSION AC084884

VERSION	AC0084884.2	GI:114327746
KEYWORDS	HTG.	
SOURCE	Oryza sativa.	
ORGANISM	Oryza sativa	
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphorbiales; Magnoliophyta; Liliopsida; Poales; Poaceae;	
REFERENCE	1 (bases 1 to 134553)	
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Henry,D. and Simmons,J.	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 134553)	
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Rambo,T., Henry,D. and Simmons,J.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-NOV-2000) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	
REFERENCE	3 (bases 1 to 134553)	
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Henry,D. and Simmons,J.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-JUN-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	
REFERENCE	4 (bases 1 to 134553)	
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Henry,D., Simmons,J., Thummond,S.K. and Mao,L.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-JUN-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	
REFERENCE	5 (bases 1 to 134553)	
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C., Kim,H., Rambo,T., Henry,D., Simmons,J., Thummond,S.K. and Mao,L.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-FEB-2002) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	
COMMENT	On Jun 7, 2001 this sequence version replaced gi:11415151. This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >30: an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone: and the assembly was confirmed by restriction digest. From 3885 to 58612, 75573 to 75592, and 79587 to 79608 are single sub-clone areas but good quality. From 107050 to 107143 there is single sub-clone coverage, but with 2 different chemistries and good quality. From 89209 to 92279 there are transposon reads but also subclones. From 90219 to 90540 there are PCR reads but also subclones.	
FEATURES	Location/Qualifiers	
SOURCE	1..134553	
	/organism="Oryza sativa"	
	/db_xref="taxon:4530"	
	/chromosome="10"	
	/clone="OSJNBa0031A07"	
	/complement(<1..1522)	
gene	/gene="OSJNBa0031A07.1"	
	/note="Contains similarity to leucine-rich repeats complementing protein"	
	/complement(join(<1..337,704..961,1220..1522))	
CDS	/gene="OSJNBa0031A07.1"	
	/note="3' Partial - at exon: 29..337"	
	/codon_start=1	
	/product="Putative leucine-rich repeats containing protein"	
	/protein_id="AA174272.1"	
	/db_xref="gi:18497053"	
	/translation="MASASGVVNALPDDLLAEFVRVAAGAKKDDSCALYRRRRGERSRAARAVPDGFDGDVAVRCADRPGLADVLDHGILTAAGSAAAEKSRQCGWENENKIDQEMQCTLSEDTQKENGSDVNPSTFDGLLHLIGCKGELKTLVNWFLHISKRGVLTANRCRNIGSLAIGGYVWHGILITLAECCNSELKLCGVDELIDNIEFLVYTRKSLVSLDISEFCNQCTIYSRLVIGTYCHNLEVLVSESKHYNEKMGIM	



RESULT 13  
AC098565 166757 bp DNA linear PLN 03-FEB-2002  
LOCUS AC098565/c  
DEFINITION Oryza sativa chromosome 10 clone OSJNBa0028C16, complete sequence.  
ACCESSION AC098565  
VERSION AC098565.3 GI:18056687  
KEYWORDS HTG.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,  
Saski,C., Henry,D., Oates,R. and Simmons,J.  
Rice Genomic Sequence  
Unpublished  
2 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,  
Saski,C., Henry,D., Oates,R. and Simmons,J.  
Direct Submission  
Submitted (24-OCT-2001) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
3 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,  
Saski,C., Henry,D., Oates,R. and Simmons,J.  
Direct Submission  
Submitted (12-NOV-2001) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
4 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,  
Saski,C., Henry,D., Oates,R., Simmons,J., Thurmond,S.K. and Sun,S.  
Direct Submission  
Submitted (22-NOV-2001) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
5 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,  
Saski,C., Henry,D., Oates,R. and Simmons,J.  
Direct Submission  
Submitted (04-JAN-2002) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
6 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,  
Saski,C., Henry,D., Oates,R., Simmons,J., Thurmond,S.K. and Sun,S.  
Direct Submission  
Submitted (30-JAN-2002) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
7 (bases 1 to 166757)  
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,  
Saski,C., Henry,D., Oates,R., Simmons,J., Thurmond,S.K. and Sun,S.  
Direct Submission  
Submitted (03-FEB-2002) Clemson University Genomics Institute,  
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA  
On Jan 4, 2002 this sequence version replaced gi:16904378.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred  
quality>30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by more than one plasmid subclone; and the assembly was  
confirmed by restriction digest. Prom base 102398-102518 there is  
some low quality. Bacterial transposons span the areas located at  
26317-28840 and 67280-69982. The assembly overlaps from base  
1-15763 with OSJNBa0031A07 (accession #AC084884). The overlap is  
from bases 118791-134553 of OSJNBa0031A07.

FEATURES  
Source  
1..166757  
/organism="Oryza sativa"  
/db\_xref="taxon:4530"  
/chromosome="10"  
/clone="OSJNBa0028C16"  
complement(3591..5736)  
repeat\_region

/note="Similar to Tourist\_O11 MITE element"  
7283..7478  
/note="Similar to Wanderer\_O11 MITE element"  
7901..7918  
gene  
/note="Simple Sequence Repeat (CTC)n"  
9577..14188  
/gene="OSJNBa0028C16.1"  
/note="Contains similarity to wall-associated protein  
kinase"  
join(9577..10489,12089..12262,13194..14188)  
/gene="OSJNBa0028C16.1"  
CDS  
/product="Putative wall-associated protein kinase"  
/codon\_start=1  
/protein\_id="AAL69419.1"  
/db\_xref="GI:18425241"  
/translation="MAFWLADLLIILASAEVAVAGAPASGARGDIDIPYGGIS  
PNCGRGFEIACNPRNDSEGMPTLAANGRTIHTOSLIVAPREKWLPAAYOCY  
SNSITDSFYGVLDLNNGTGYRISDSRNNEVYIGNTISTYTGNSGGGPGYAGLYT  
GCVSYCNDSSSARDSSACAGCCCHIDISPLSDNVVSVFGPMKRGQVDSPODSEFLV  
DKNEYERSADLMKDLNRPVWLDMAIRDSVTCPLPEQEKRPAGYACMSDNSECVN  
STNGPGYCKKQKQGDGNPYVDQCCXIDINEDVSNKKKPCYGCVCNNIPEDYECN  
RVGOWSGEGEPKQECSSAKPPLAARALGITTGSEFLIYAVLETLMMKNNIPEDEYEFK  
KNGSVLQKVDNKTIFTKDELKRTLNNSVEYVGGSEGVYGTLEDNPPVAKNSIE  
VNAARKDDPNEVITIOSMHHNNITKLCGCELVDPMLYEPAPAKGNQDILKRDAN  
IPLEGLRDIAIESAGLRYMHSSTRTIRHODVPAWILTDKFIPTISDGTSLK  
LVNDKDTMPEVNSKGYIDVFRTGHTLQKSDVVSFVGLLELCRKEPTIYGENCSL  
IIEFONAYPDENSGRIMFDEINBEDDILIEIGRLAMECLKEKEVERPDMKEVAER  
FVMLRSRRCG"  
complement(20437..20586)  
repeat\_region  
/note="Similar to Explorer\_Os1 MITE-like element"  
2230..25021  
/gene="OSJNBa0028C16.2"  
/note="Contains similarity to gag-pol polyprotein"  
22230..25021  
/note="Contains similarity to gag-pol polyprotein"  
join(22230..22391,22449..22860,23130..25021)  
/gene="OSJNBa0028C16.2"  
CDS  
/codon\_start=1  
/product="Putative gag-pol polyprotein"  
/protein\_id="AAL69420.1"  
/db\_xref="GI:18425242"  
/translation="MAEKPPPSRSGFYKDKMLKGLLDVNEGNNVPIDPKFTPEO  
KKEFEALOOKKIKYVADDPGTSFKDDSGKOAPGSGASPTIKGATDFLQNDGNS  
QGVHGVQGVGVHGVQDGAQGLQGNLNNONSVADQFPNNFDDRYDAVHALLINQGS  
VLVNSLSNMKSIADSIADSIADSIADSIADSIADSIADSIADSIADSIADSIADSI  
PPHQOPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPLQPL  
VQPGGVVMPQVDEIHLVNTQPLHNYGQGNLNYQVPSPOAOLALADOPASO  
YQGPPEMOOSQGAQOQPYANVITADVMEQGLKREKESLYRQPIPEMFERVPLN  
WEVLPDSKFSQDGVSTIEHLSRFLAQGSEASVDALVRLPSTLGSATWFSLSL  
LVGSINSMADLERQGFHSYFYGGHEMKLSDLAIKQRHDEPQVETIQRERNKCFIS  
LSLTDALQDLAEQGMIAPIREKFSSEDESDLSOLIKLLEHRSAPRANKSKKLVNH  
VCLVNYGSDDEDESEIATAAEWRSKVIIPCOMVSSGKEEYDDIDIRANKIFPLL  
REKQIDLPNGHTIPSAEILGKKRYCKWHNSRSHSTGDFVQFOQIOVAIEGKIRFDY  
SKPKMAYDNPVNVNVTAGQIADIVRARGQANSATKIDYTHHEYEDGDPNPGCE  
FFRCNBSGRRLPSIENCPCSDIYESSRSYTRNRLRQTVSHVQRGPNQDHGO  
EDVNIKKNQCBPGIFTKNQKN"  
complement(26926..27117)  
/gene="OSJNBa0028C16.3"  
/note="Hypothetical protein"  
complement(26926..27117)  
/gene="OSJNBa0028C16.3"  
CDS  
/product="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AAL69421.1"  
/db\_xref="GI:18425243"  
/translation="MASSPARGGKRPTRALDGEVGVARGLANPPTATYARVLCGSS  
GDRRPRKAKAAAGGARG"  
complement(28733..30643)  
/gene="OSJNBa0028C16.4"  
/note="Contains similarity to gag-pol polyprotein"  
complement(28733..30643)  
/note="Contains similarity to gag-pol polyprotein"  
complement(join(28733..29538,29893..30643))  
repeat\_region  
CDS



Center lone name: CH230-206C20	
----- Summary Statistics -----	
Assembly program: Phrap, version 0.990329First call to findPhrapList	
Consensus quality: 141246 bases at least Q40	
Consensus quality: 147079 bases at least Q30	
Consensus quality: 151264 bases at least Q20	
Estimated insert size: 144279; sum-of-contigs estimation	
Quality coverage: 0x in Q20 bases; agatose-fp estimation	
Quality coverage: 2.3x in Q20 bases; sum-of-contigs estimation	
-----	
* NOTE: Estimated insert size may differ from sequence length	
* (see <a href="http://www.hbsc.bcm.tmc.edu/docs/genbank_draft_data.html">http://www.hbsc.bcm.tmc.edu/docs/genbank_draft_data.html</a> )	
* NOTE: This is a "working draft" sequence. It currently	
* consists of 56 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
-----	
1	9077: contig of 9077 bp in length
9078	9177: gap of unknown length
9178	17229: contig of 8052 bp in length
17330	17329: gap of unknown length
17330	22707: contig of 5378 bp in length
22708	22807: gap of unknown length
22808	29002: contig of 6195 bp in length
29003	29102: gap of unknown length
29103	35078: contig of 5976 bp in length
35079	35178: gap of unknown length
35179	40876: contig of 5698 bp in length
40877	40976: gap of unknown length
40977	47048: contig of 6072 bp in length
47049	47148: gap of unknown length
47149	52265: contig of 5117 bp in length
52266	52365: gap of unknown length
52366	56605: contig of 4241 bp in length
56607	56706: gap of unknown length
56707	60880: contig of 4174 bp in length
60881	60980: gap of unknown length
60981	65627: contig of 4647 bp in length
65628	65727: gap of unknown length
65728	69756: contig of 4029 bp in length
69757	69856: gap of unknown length
69857	72872: contig of 3016 bp in length
72873	72972: gap of unknown length
72973	76267: contig of 3295 bp in length
76268	76367: gap of unknown length
76368	78769: contig of 2402 bp in length
78770	78869: gap of unknown length
78870	82079: contig of 3210 bp in length
82080	82179: gap of unknown length
82180	85534: contig of 3355 bp in length
85535	85634: gap of unknown length
85635	88442: contig of 2808 bp in length
88443	88542: gap of unknown length
88543	91631: contig of 3089 bp in length
91632	91731: gap of unknown length
91732	94958: contig of 3227 bp in length
94959	95058: gap of unknown length
95059	97723: contig of 2665 bp in length
97724	97823: gap of unknown length
97824	100565: contig of 2742 bp in length
100566	100665: gap of unknown length
100666	102726: contig of 2061 bp in length
102727	102826: gap of unknown length
102827	104699: contig of 1873 bp in length
104700	104799: gap of unknown length
104800	106919: gap of unknown length
106920	107019: contig of 2120 bp in length
107020	108912: gap of unknown length
108912	108912: contig of 1893 bp in length
108913	109012: gap of unknown length

Search completed: May 7, 2002, 15:55:53  
Job time: 14774 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.,

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 11:49:39 ; Search time 247.46 Seconds  
(without alignments)  
7888.676 Million cell updates/sec

Title: US-09-579-383-1

Perfect score: 1137

Sequence: 1 atgaatttaccgtaataa.....tcgttgaagcctgcactaa 1137

Scoring table: IDENTITY\_NUC

Searched: 1736436 seqs, 858457221 residues

3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1137	100.0	1137	22 AAC89668	P. falciparum chit
2	104.2	9.2	1764	22 AAC89669	P. gallinaceum chl
3	104.2	9.2	2500	22 AAC89670	P. gallinaceum chl
4	38	3.3	4590	22 AAH24065	Yeast AOD9604-asso
5	36.4	3.2	2258	23 ABL09857	Drosophila melanog
6	36.4	3.2	4553	23 ABL09856	Drosophila melanog
7	36.4	3.2	6488	23 ABL10302	Drosophila melanog
8	35.2	3.1	3294	23 AAS73382	DNA encoding novel
9	35.2	3.1	3294	23 AAS89403	DNA encoding novel

C	10	35.2	3.1	3294	23 AAS94292	DNA encoding novel
C	11	35.2	3.1	3296	23 AAS92034	DNA encoding novel
C	12	34.8	3.1	2111	19 AAX14467	H. pylori GPO 103
C	13	34.8	3.1	2307	23 ABL23687	Drosophila melanog
C	14	34.8	3.1	5479	23 ABL23686	Drosophila melanog
C	15	34.8	3.1	27425	21 AAS53978	DMP53 tumour suppr
C	16	34.6	3.0	4729	11 AAQ04043	Recombinant DNA fo
C	17	34.4	3.0	42000	21 AAB63349	Streptomyces globi
C	18	34.4	3.0	63164	21 AAB63348	Streptomyces globi
C	19	34.4	3.0	3172	22 AAS34789	CDNA encoding nove
C	20	34	3.0	3202	18 AAT47520	Human hepatoma-der
C	21	34	3.0	3256	22 AAS33262	DNA encoding human
C	22	34	3.0	3256	22 AAS34946	CDNA encoding nove
C	23	34	3.0	3316	19 AAV44859	Clone A073.3 codin
C	24	34	3.0	3316	22 AAF98456	Human cDNA clone A
C	25	34	3.0	3707	22 AAS34889	CDNA encoding nove
C	26	33.8	3.0	8196	23 ABL02886	Drosophila melanog
C	27	33.8	3.0	133894	17 AAT13635	ACNPV genomic DNA
C	28	33.8	3.0	1038602	20 AAT01425	Complete genome se
C	29	33.6	3.0	92407	22 AAF28549	Genomic fragment #
C	30	33.4	2.9	2100	23 ABL20941	Drosophila melanog
C	31	33.4	2.9	2149	18 AAT91115	Absidia blakesleea
C	32	33.4	2.9	2149	19 AAV59380	Nucleotide sequenc
C	33	33.4	2.9	4266	20 AAZ31004	Staphylococcus aur
C	34	33.4	2.9	4308	22 AAF54734	Nucleotide sequenc
C	35	33.4	2.9	4308	23 AAS51929	Staphylococcus aur
C	36	33.4	2.9	4329	23 ABL20940	Drosophila melanog
C	37	33.4	2.9	4329	23 AAS54586	Staphylococcus aur
C	38	33.4	2.9	5040	21 AAB11351	Rice blast fungus-
C	39	33.4	2.9	22243	18 AAV74475	Staphylococcus aur
C	40	33.2	2.9	817	22 AAH07130	Human cDNA clone (
C	41	33.2	2.9	1148	22 AAH14736	Human cDNA sequenc
C	42	33.2	2.9	1177	21 AAC77080	Human ORFX ORF2635
C	43	33.2	2.9	3102	22 AAH90796	CPE 100 coding seg
C	44	33	2.9	361	14 AAO59777	Human brain Expres
C	45	33	2.9	1539	23 AAS84960	DNA encoding novel

#### ALIGNMENTS

RESULT 1	
ID AAC89668	standard; DNA; 1137 BP.
XX AAC89668;	
XX	
DT 13-MAR-2001	(first entry)
XX	
DE P. falciparum chitinase coding sequence SEQ ID NO: 1.	
XX	
KW Malaria; mosquito; chitinase; fungal disease; parasitic disease;	
KW veterinary disease; arthropod pest; ds.	
XX	
OS Plasmodium falciparum.	
XX	
PN W0200073488-A1.	
XX	
PD 07-DEC-2000.	
XX	
PF 26-MAY-2000; 2000MO-US14536.	
XX	
PR 28-MAY-1999; 99US-0136508.	
PR 03-FEB-2000; 2000US-0180051.	
XX	
PA (TEXA ) UNIV TEXAS SYSTEM.	
XX	
PI Vinetz JM;	
XX	
DR WPI: 2001-061553/07.	
DR P-PSDB; AAB49985.	
XX	
PT New nucleic acid encoding a Plasmodium species chitinase is useful for	

PT preventing transmission of malaria by mosquito feeding on subject that  
 PT may harbor Plasmodium species organisms -  
 XX  
 PS Claim 4: Page 123: 137pp: English.

CC The present invention provides the protein and coding sequences of the  
CC *Plasmodium falciparum* and *P. gallinaceum* chitinase enzymes. These  
CC organisms are the cause of malaria in humans. The sequences are useful  
CC in the prevention and treatment of malaria, fungal diseases, parasitic  
CC diseases and veterinary diseases, in preventing the transmission of  
CC malaria and in the control of arthropod pests in agriculture.

Sequence 1137 BP; 321 A; 263 C; 257 G; 296 T; 0 other;

Query Match	100.0 %	Score 1137	DB 22	Length 1137
Best Local Similarity	100.0 %	Pred. No. 0		
Matches 1137	0	Mismatches	0	Gaps 0

Qy 1 atgaatttcgcgtaaaatactctcttcctggtattgtctgctgtgcctgtcct 60  
|||||  
Db 1 atgaatttcgcgtaaaatactctcttcctggtattgtctgctgtgcctgtcct 60

Qy	Db
61 actatgtttctgttatcgaagtcacgltgcgtccgggtgaatccgcaaaaaccg 120	61 actatgtttctgttatcgaagtcacgltgcgtccgggtgaatccgcaaaaaccg 120

Qy 121 cgtgaattatcaaaaagctttaagaatccggtgaaggtcatccagggttactatccg 180  
|||||  
Db 121 cgtgaattatcaaaaagctttaagaatccggtgaaggtcatccagggttactatccg 180

Qy 181 tccctggtgagctataatcacaactgaaagacctgaaccggaacctgaagtggtccac 2400  
|||||  
Db 181 tccctggtgagctataatcacaactgaaagacctgaaccggaacctgaagtggtccac 2400

**QY**    241 atgagcttctgttaaaatgatatttagcctacgacctccatctgaatcatcgtagtagccg 3000  
         |||||  
**Db**    241 atgagcttctgttaaatgatatttagcctacgacctccatctgaatcatcgtagtagccg 3000

QY 301 ctgctgctttaaagccgcatgttgccctgagtatatttgcttgaagaggtacttcaatgat 3600  
|||||  
|||  
Db 301 ctgctgctttaaagccgcatgttgccctgagtatatttgcttgaagaggtacttcaatgat 3600

**QY** 361 gccatgaaatcgcgaagctcgccggacatttatgtctgtagccctggcgctgag 420  
**Db** 361 gccatgaaatcgcgaagctcgccggacatttatgtctgtagccctggcgctgag 420

QY 421 acctaacatcgcgtccctccttgattctgcactcaacgcygttgaaaaatcgcaarctg 480  
|||||  
db 421 acctaacatcgcgtccctccttgattctgcactcaacgcygttgaaaaatcgcaarctg 480

QY 481 gtggaatgaactggcgcttcgatgtatgatgtgaatgaacggaacgagcttggat 540  
|||||  
Db 481 gtggaatgaactggcgcttcgatgtatgatgtgaatgaacggaacgagcttggat 540

**QY** 541 ggtctgtaatgataagaanaagcgcactcttctgtacaatacgtagccaactgcggaa 600  
|||||  
**Db** 541 ggtctgtaatgataagaanaagcgcactcttctgtacaatacgtagccaactgcggaa 600

```

601 tacactgtgtagataaacatgcagcatcagccagtcctctaagcgcgacctgaagctgc 660
      |||
db    601 taacatgtgtgatataaacatgcagcatcagccagtcctctaagcgcgacctgaagctgc 660

```

**Dd**

```
661 atcggttcaacgaccggaaaaatctgtatgatgcgaagctccgtatacagcaaa 720
```

Db 721 tatttcaacaacgcgacgtttaagaagaactgttaagcgcagccagatgcatctgcg 780

Db	761	ggggtgcccattcactccggtatgnaaanaacctgaaagatatgattgacatggtgtgtgttcgcg	840
Qy	841	acgttcaactacacgaaactctaccgattcttaccggtatgnaaanaatatlacagctctac	900
Db	841	acgttcaactacacgaaactctaccgattcttaccggtatgnaaanaatatlacagctctac	900
Qy	901	gcatactctgtgtaaaaggtacgatttaagtgatcatatggtgttcccttgattgttcgcg	960
Db	901	gcatactctgtgtaaaaggtacgatttaagtgatcatatggtgttcccttgattgttcgcg	960
Qy	961	tccaagcggttccaaccccgaaacgataaatctgtgtaaatctctcggcgattctgtaaa	1020
Db	961	tccaagcggttccaaccccgaaacgataaatctgtgtaaatctctcggcgattctgtaaa	1020
Qy	1021	actgaaaaacaactgtaataacgcgcagatggtttcgcgctgtggtctctgtgccagcgac	1080
Db	1021	actgaaaaacaactgtaataacgcgcagatggtttcgcgctgtggtctctgtgccagcgac	1080
Qy	1081	aacgcggcccaatatgaaacagctcggcgatcgatgtaactcgtttgaaagcctgcactaa	1137
Db	1081	aacgcggcccaatatgaaacagctcggcgatcgatgtaactcgtttgaaagcctgcactaa	1137

RESULT	2
AAC89669	
ID	AAC89669 standard; DNA; 1764 BP.

AC	AAC89669;
XX	
DT	13-MAR-2001 (first entry)

DE	P. gallinaceum chitinase coding sequence SEQ ID NO: 2.
XX	
KW	Malaria; mosquito; chitinase; fungal disease; parasit

XX  
OS  
XX  
Plasmodium gallinaceum

XX 07-DEC-2000.  
PD  
XX

XX	28-MAY-1999;	99US-0136508.
PR	03-FEB-2000;	2000US-0180051.
PR		

PA (TEXA ) UNIV TEXAS SYSTEM.  
XX  
PI Vinetz JM;

DR WPI; 2001-061553/07  
DR P-PSDB; AAB49986.  
XX

PT	preventing transmission of malaria by mosquito feeding on subject that
PT	may harbor <i>Plasmodium</i> species organisms -
XX	

XX The present invention provides the pro  
CC Plasmodium falciparum and P. gallinace  
CC

CC in the prevention and treatment of malaria, fungal diseases, parasitic  
CC diseases and veterinary diseases, in preventing the transmission of  
CC malaria and in the control of arthropod pests in agriculture.

Sequence 1764 BP; 717 A; 227 C; 294 G; 526 T; 0 other;

Query Match	9.28;	Score 104.2;	DB 22;	Length 1764;
Best Local Similarity	48.08;	Pred. No. 1.2e-21;		
Matches 468; Conservative	0;	Mismatches 483;	Indels 24;	Gaps 5;

```

QY 110 gcaaaacccgctgaataatcatcaaaacglttaagaatcggtaagatcatccagg 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 gaaatcaccaagacaatlttagagagatataaaaaagaaacaggtattatagcag 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 gttactatccgctcggtgtgactataacac-----acctgaagcctcggaac 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 gatactatggtctacaggaacatgtagtagagcaaaacatgtagtttcaaac 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 cgaacctgaacgtggttcacatgagctttgctaaatgagtttaagctacagctccatgt 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 caatggtgtcaatttatatatattgcatcttgcgcataatattgttatatgttcatcta 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 aatccatcgttagtagcccgctgctgtttaaaagcctgattggcctgtagtatattgttc 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 gaccatttaatgagaacaaagattccatttaagaanaacagcgttaagatatgaac 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 tgaatgagtaactactatgatgcatgaatctgacgaagctgctcggaacattatagc 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 atgtagatgctcttaagaaattagacgtatcagaanaagtagccagatgtaattatc 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 tgcctgagctggcgctgtagacatccatccgtccctcctcgatcctgacacagcg 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 ttctatccttaggttggaagaacct--atagatagatatagaanaagaaattgattag 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 ttgaaaaaatcgcaaatctgtgtgtagaacttgggtctcgatgtagtatgtagattacg 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 623 tggataaataatgaaagctgtgtaattgattagattgtagtggtagattgtcgtg 682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 521 aacggaaagcgagctttagtgcgtgaatgataaagaanaagcgactcttctgacaat 580
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 aaccacatgggagatttcaacaacttaaatgaattaaatttcaattattattataat 742
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 581 acgtacacaactgvcgcaatcatgltgatgataaactgacatcgacacgagctc 640
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 taattaaactgttgaagaanaactatccggaagaagaatttaattcatttggttcat 802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ctaatgvcgtctgagctgcatcggttcaacgacccgaaaaaaatctgtatgtagtagc 700
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 caaatgctgcatlatacagctgttcagagatgcatc-----ctctgtaagatgtag 856
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 701 aagctccgtataacgcaaatattt---caacaaccgagcgttaagaagaactgttac 757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 aatcccatataacacataatttctgtcgaacaataagaaacataaagaattacata 916
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 758 gcgacgccaagatgvcgcatcgcggtggtgcccactactacgtgtagacaacccgaaagata 817
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 917 gggcgcagcgatgttatacagcaggaactttatataatttttaataagcaagaagata 976
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 818 tgattgacatggtgttggtagagacgttcaactacacgaactctacgattacagttta 877
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 977 aaatagatcctgtatttatttcaacatacaatt---tagaacctcaaatccagatata 1033
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 878 tgaagaattatacagacccctacacgatactatgtagtaaaaaagtagattagcgtatcatca 937
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1034 tggtagatagtagtactatcccatattatatttgggtttaaatatatacacatacatat 1093
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 938 tgggcttacccctgtagtcccgatccagcgcgttcaccccgagagataaaatgycgtgtaa 997
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1094 taggttttcaattgaacataacagaggtgtagtttagtcccgaaataaagaattatag 1153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 998 aatcatcgcgatttctglaaanaactgaanaacaaactgataaagcgagatggtttc 1057
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1154 aattgtagtaaaaaacatacatgataaatacaaaataataatagggcagatggtatag 1213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1058 gccgtgtgtcctgt 1072
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1214 ggtatggtcattat 1228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3  
AAC89670  
ID AAC89670 standard; DNA; 2500 BP.

```

XX AAC89670;
AC 13-MAR-2001 (first entry)
DT
XX
XX
DE P. gallinaceum chitinase coding sequence SEQ ID NO: 5.
XX
XX KW Malaria; mosquito; chitinase; fungal disease; parasitic disease;
XX veterinary disease; arthropod pest; ds.
OS Plasmidium gallinaceum.
XX WO200073488-A1.
XX
XX PD 07-DEC-2000.
XX
XX PF 26-MAY-2000; 2000MO-US14536.
XX
XX PR 28-MAY-1999; 99US-0136508.
XX PR 03-FEB-2000; 2000US-0180051.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX PA Vinetz JM;
XX
XX PI WPI; 2001-061553/07.
XX
XX DR
XX
XX PT New nucleic acid encoding a Plasmidium species chitinase is useful for
XX preventing transmission of malaria by mosquito feeding on subject that
XX may harbor Plasmidium species organisms
XX
XX PS Disclosure; Page 127-128; 137pp; English.
XX
XX
XX CC The present invention provides the protein and coding sequences of the
XX Plasmidium falciparum and P. gallinaceum chitinase enzymes. These
XX organisms are the cause of malaria in humans. The sequences are useful
XX in the prevention and treatment of malaria, fungal diseases, parasitic
XX diseases and veterinary diseases, in preventing the transmission of
XX malaria and in the control of arthropod pests in agriculture.
XX
XX SQ Sequence 2500 BP; 993 A; 283 C; 365 G; 858 T; 1 other;

Query Match 9.2%; Score 104.2; DB 22; Length 2500;
Best Local Similarity 48.0%; Pred. No. 1.4e-21;
Matches 468; Conservative 0; Mismatches 483; Indels 24; Gaps 5;

QY 110 gcaaaacccgctgaataatcaaaacglttaagaatcggtagaagatcatccagg 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 gaaatcaccaagacaatlttagagagatataaaaaagaaacaggtattatagcag 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 gttactatccgctcggtgtgactataacac-----acctgaagcctcggaac 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 gatactatggtctacaggaacatgtagtagagcaaaacatagattgtattcaaac 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 cgaacctgaacgtggttcacatgagctttgctaaatgagtttaagctacagctccatgt 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 caatggtgtcaatttatatatattgcatcttgcgcataatattgtatatgttcatcta 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 aatccatcgttagtagcccgctgctgtttaaaagcctgattggcctgtagtatattgttc 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 gaccatttaatgagaacaaagattccatttaagaanaacggttagaattgaac 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 tgaatgagtaactactatgatgcatgaatctgacgaagctgctcggaacattatagc 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 atgtagatgctcttaagaaattagacgtatcagaanaagtagccagatgtaattatc 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 tgcctgagctggcgctgtagacatccatccgtccctcctcgatcctgacacagcg 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 ttctatccttaggttggaagaacct--atagatagatatagaanaagaaattgattag 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 ttgaaaaaatcgcaaatctgtgtgtagaactggtcgtcgttagtattgttagattag 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 898 tggataaataatgaagcttgtaattgattagatggtgtagatattgactg 957
QY 521 aaccgaacgagcttctgattgctgtaataagaacgagctcttctgacat 580
Db 958 aaccacatgggaatttaccacttaataatttcaattatattatataat 1017
QY 581 acgtgaccaactcgagacatcatggtgatgataactgacgacgacgctct 640
Db 1018 taattactctgttaagaanaactcttcggaagaagaattcaattctgttct 1077
QY 641 ctaatggcgtctgagctgcatcggttcaacgaccggaanaaaactgtagatgacg 700
Db 1078 caatgctcattatcatcgcttccagagattgcatc-----cttcgtaagaag 1131
QY 701 aagctcgtaataacgacgaatatt---caacaacgagcttagaagaagactgtac 757
Db 1132 aatcccatataacacataaatttctgctgaacaataagaacaataagaattacata 1191
QY 758 ggcgagcccaagatgcatctgagggtggcgcatctacgtgagacaacctgaaagata 817
Db 1192 gggcagcagcagctgtatcacgacgaactttatattatatttatacagcaagaga 1251
QY 818 tgattgacatgggtgttctgacagacgttcaactacacgaactctcagattctaccgta 877
Db 1252 aatatagatctgtatttatttcaacatatacatt---tagaaactacaattccagatalaa 1308
QY 878 tgaagaattatacgaactcctcagcatatctgtgtaaaagtagattacgtagatcata 937
Db 1309 tggtagatatgacttattccattatatttggtttaataatacatcacacaacatc 1368
QY 938 tgggtctcacccgtatgttccgctccagcgcgttcaaccggaacgaataaactcgtgtaa 997
Db 1369 tagtttcttcatgagaacatacagagggtgatttagtcccgaaataaagaattatag 1428
QY 998 aatctatcgcgagatttctgtaaaactgaaacaaactgaataacgacgagtggttcg 1057
Db 1429 aattgtagtagaataacatacatgataaaatcaaaataataatagggcagatggtatag 1488
QY 1058 gccctgtgctcgtg 1072
Db 1489 ggatagtgcatattat 1503

```

```

FT misc_feature 3819
FT /tag= e
FT /note= "Represented as * in the specification"
FT misc_feature 3862
FT /tag= f
FT /note= "Represented as * in the specification"
FT misc_feature 3864
FT /tag= g
FT /note= "Represented as * in the specification"
FT misc_feature 3888
FT /tag= h
FT /note= "Represented as * in the specification"
FT misc_feature 3890
FT /tag= i
FT /note= "Represented as * in the specification"
FT misc_feature 3912
FT /tag= j
FT /note= "Represented as * in the specification"
FT misc_feature 3914
FT /tag= k
FT /note= "Represented as * in the specification"
FT misc_feature 3938
FT /tag= l
FT /note= "Represented as * in the specification"
FT misc_feature 3939
FT /tag= m
FT /note= "Represented as * in the specification"
FT misc_feature 3941
FT /tag= o
FT /note= "Represented as * in the specification"
FT misc_feature 3943
FT /tag= p
FT /note= "Represented as * in the specification"
FT misc_feature 4361
FT /tag= q
FT /note= "Represented as * in the specification"

```

WO200133977-A1.

17-MAY-2001.

06-NOV-2000; 2000WO-A001362.

05-NOV-1999; 99AU-0003875.

(META-) METABOLIC PHARM LTD.

Belyea CI, Ng FM, Vaughan P;

WPI; 2001-328876/34.

New organisms containing nucleic acid encoding a growth hormone fragment which modulates lipid metabolism are useful to produce dietary aids for obesity and in the meat production industry.

Disclosure; Page 48-50; 54pp; English.

The invention relates to novel transgenic organisms useful in the production of functional food and drink products for the treatment or prevention of obesity via the regulation of lipid metabolism. The organisms comprise a polynucleotide encoding a growth hormone fragment capable of stimulating the activity of hormone-sensitive lipase (the key enzyme in lipolysis) and inhibiting acetyl CoA carboxylase (the key enzyme in lipogenesis). The growth hormone fragment preferably contains at least the disulphide-bonded loop of a mammalian growth hormone (but is not the full-length growth hormone) and is optionally linked to an epitope tag or heterologous fusion protein partner. The transgenic organism may be a microorganism used to produce a fermented product (e.g., yeast), or an edible plant or animal or cell thereof. Food or drink made using methods of the invention are used to modify fat/lean ratio, lipid metabolism or food use in a mammal. In particular, the food or drink products may be used to treat or prevent obesity, particularly in humans, and may also be used to improve the fat/lean ration of



XX Claim 1; SEQ ID NO 24050; 21pp + Sequence Listing; English.  
PS  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB57737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 4553 BP; 1404 A; 993 C; 901 G; 1255 T; 0 other;

Query Match 3.2%; Score 36.4; DB 23; Length 4553;  
Best Local Similarity 51.2%; Pred. No. 2.5;  
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY 502 ggtatgtatgtagattacgaacgacgacgacgttgatgtgtcgaatgaataaagaaaa 561  
DB 3150 GGAACATCATGCTGATGCTTAGAGAACGTGATGTGTCTTCTTGCTAAAGTAATCA 3091  
OY 562 gcggactctcttgatcaatacgtgacacactgcgcgaatacatgtgtgatgataactg 621  
DB 3090 ACAGATTTCGTTAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3031  
OY 622 atcagcatcagccagctcctcctaatacgcgctcgtgagctgacatcggtt 667  
DB 3030 ATAAACATAGACTTTCTCAAACTTAACAACAATGCAATCGGCT 2985

RESULT 7  
AB10302/c  
ID AB10302 standard; cDNA; 6488 BP.  
XX  
XX AB10302;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 25388.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PF  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-0614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX P-PSDB; AB566199.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Claim 1; SEQ ID NO 25388; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB57737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 6488 BP; 1991 A; 1324 C; 1311 G; 1862 T; 0 other;

Query Match 3.2%; Score 36.4; DB 23; Length 6488;  
Best Local Similarity 51.2%; Pred. No. 3;  
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY 502 ggtatgtatgtagattacgaacgacgacgacgttgatgtgtcgaatgaataaagaaaa 561  
DB 393 GGAACATCATGCTGATGCTTAGAGAACGTGATGTGTCTTCTTGCTAAAGTAATCA 334  
OY 562 gcggactctcttgatcaatacgtgacacactgcgcgaatacatgtgtgatgataactg 621  
DB 333 ACAGATTTCGTTAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274  
OY 622 atcagcatcagccagctcctcctaatacgcgctcgtgagctgacatcggtt 667  
DB 273 ATAAACATAGACTTTCTCAAACTTAACAACAATGCAATCGGCT 228

RESULT 8  
AAS73382  
ID AAS73382 standard; cDNA; 3294 BP.  
XX  
XX AAS73382;  
AC  
XX  
XX 13-FEB-2002 (first entry)  
DT  
XX  
XX DNA encoding novel human diagnostic protein #9186.  
DE  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200175067-A2.  
PN  
XX  
XX 11-OCT-2001.  
PD  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
PF  
XX  
XX 31-MAR-2000; 2000US-0540217.  
PR  
XX  
XX 23-AUG-2000; 2000US-0649167.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX  
XX WPI; 2001-639362/73.  
DR  
XX  
XX P-PSDB; ABG09195.  
DR  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 1; SEQ ID NO 9186; 103pp; English.  
PS  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The

polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques for identifying expressed genes. (II) is useful to treat disease states involving restoring normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS6197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPRO at [ftp.wipro.int/pub/published\\_pcl\\_sequences](http://wipro.int/pub/published_pcl_sequences).

	Query Match	3.1%	Score 35.2	DB 23	Length 3294
	Best Local Similarity	52.0%	Prod. No. 4.9		
	Matches 79	Conservative 0	Mismatches 73	Indels 0	Gaps 0
QY	549	tgataaagaaaagagagactctttgttacaatgaacggaacacgtgcgaatatactg	608		
DB	2202	tgttttggaataaacacatcttcgaatcgtggaacttttgaagaccactttttgacgcggtat	2261		
QY	609	tgtatataactgtaacagcatcagccagctctctctaattgcygcgtcgtgaatcgatcggttt	668		
DB	2262	tgcagataaacgcgatacgtcttcattgatatatcgcaataaagacgtttgccttcgtgacct	2321		
QY	669	caacgaccccgaaaaaatctgtatgtagacg	700		
DB	2322	caacgacacglttgataataatgttcgctaatt	2353		

RESULT	9
AA589403	
ID	AA589403 standard; cDNA; 3294 BP.
XX	
AC	
XX	AA589403;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #25207.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	Human; supplement; medical imaging; diagnostic; genetic disorder; ss
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
PD	
XX	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
XX	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PL	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG25216.
XX	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -

XX Claim 1: SEQ ID No 25207; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and genome mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX

XX Sequence 3294 BP; 720 A; 926 C; 943 G; 705 T; 0 other;

Query Match	3.1%	Score 35.2	DB 23	Length 3294
Best Local Similarity	52.0%	Pred. No. 4.9		
Matches 79	Conservative	0	Mismatches 73	Indels 0
				Gaps 0
QY 549	tgtataagaanaagcgagactctctgttacaatcagcttgaccaacatgcgcgaatacatgtr	608		
DB 2202	tgtttatgtaagaataaacacatcttgatgvgtgactttgagaacacatttttgaccggttat	2261		
QY 609	tgtatgataaacatgcatcagcatcaagccagtcctctataatggcgcctctgagctgcatcggtt	668		
DB 2262	tgcagataaacccgatcagcttcattgatatacagaataaagatttgccttcgtgacct	2321		
QY 669	caacgaccccgaaaaaatctgtatgtagacg	700		
DB 2322	caacgacacagtttgataataatgctgcctaatg	2353		

RESULT	10	
AA594292/c		
ID	AA594292	standard; cDNA; 3294 BP.
XX		
AC	AA594292;	
XX		
DT	13-FEB-2002	(first entry)
XX		
DE	DNA encoding novel human diagnostic protein #30096.	
XX		
KN	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX	food supplement; medical imaging; diagnostic; genetic disorder; ss	
XX		
OS	Homo sapiens.	
XX		
PN	WO200175067-A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US08631.	
XX		
PR	31-MAR-2000; 2000US-0540217.	
PR	23-AUG-2000; 2000US-0649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Drmanac RT, Liu C, Tang YT,	
XX		

DR WPI, 2001-639362/73.  
DR P-PSDB; ABG30105.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 1; SEQ ID No 30096; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 3294 BP; 705 A; 943 C; 926 G; 720 T; 0 other;

Query Match 3.1%; Score 35.2; DB 23; Length 3294;  
Best Local Similarity 52.0%; Pred. No. 4.9;  
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 549 tgataaagaagaacgagactcttctgtacatacgtgacaaactgcgcgaatacatgtg 608  
DB 1093 TGTATTGGAATACCACTTCCTTCATGGTGACTTTGAGAACCACTTTTGGACCGCGTTAT 1034  
QY 609 tgatgataaactgatacgcacgcagtcctcctaattgycgctcgtgacgtcagttt 668  
DB 1033 TGCAGATTAACCGATACGCTCATGATGATGCAATACGATTGGCCCTTCGTGACCT 974  
QY 669 caacgacccgaaaaaactcgtatgtagtgcg 700  
DB 973 CAACGAAACAGTTGGATTAATATGTCGCTAATG 942

RESULT 11  
AAS92034  
ID AAS92034 standard; cDNA; 3296 BP.  
XX  
AC AAS92034;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #27838.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensics;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2001/5067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001MO-US08631.  
XX

PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI, 2001-639362/73.  
DR P-PSDB; ABG27847.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 1; SEQ ID No 27838; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 3296 BP; 720 A; 927 C; 940 G; 709 T; 0 other;

Query Match 3.1%; Score 35.2; DB 23; Length 3296;  
Best Local Similarity 52.0%; Pred. No. 4.9;  
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 549 tgataaagaagaacgagactcttctgtacatacgtgacaaactgcgcgaatacatgtg 608  
DB 2202 tgtatggaataaaccatcttcgattgtagtgcagacacatttttgacgcggtat 2261  
QY 609 tgatgataaactgatacgcacgcagtcctcctaattgycgctcgtgacgtcagttt 668  
DB 2262 tgcagataaaccgatacgtctcattgataatgcgcaaaacgaattgcccctcgtgac 2321  
QY 669 caacgacccgaaaaaactcgtatgtagtgcg 700  
DB 2322 caacgacagttgataataatgtagtgcgta 2353

RESULT 12  
AAX14467/C  
ID AAX14467 standard; DNA; 2111 BP.  
XX  
AC AAX14467;  
XX  
DT 31-MAR-1999 (first entry)  
XX  
DE H. pylori GHPO 1039 gene.  
XX  
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
KW peptic ulcer disease; ss.  
XX  
OS Helicobacter pylori.  
XX



```

XX XX Key Location/Qualifiers
FH FH 31..2070
FT CDS /*tag= a
XX
XX MO9843478-A1.
XX
XX 08-OCT-1998.
XX
XX 01-APR-1998; 98WO-US06371.
XX
XX 29-JUL-1997; 97US-0902615.
PR 01-APR-1997; 97US-0833457.
PR 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
PI
XX WPI: 1998-542293/46.
DR
XX P-PSDB; AAW98748.
XX
XX Claim 1: Page 1597-1600; 2054pp; English.
XX
XX This sequence represents a polynucleotide of the invention. It was
CC isolated from Helicobacter pylori and encodes a H.pylori GHP0 protein.
CC The polypeptides can be used for preventing or treating Helicobacter
CC infections, and gastroduodenal diseases associated with these
CC infections, including acute, chronic, and atrophic gastritis, and peptic
CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
XX Sequence 2111 BP; 752 A; 392 C; 426 G; 541 T; 0 other;
XX
XX Query Match 3.1%; Score 34.8; DB 19; Length 2111;
XX Best Local Similarity 58.8%; Pred. NO. 5.1;
XX Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
XX
XX QY 236 ttcaatgagcttgcttaaatgattgaagctacgactccatgcatcgtaggtta 295
XX ||||| | | | | | | | | | | | | | | | | | | | | | | |
XX DB 1889 TTCAAAACATGATGGCCATATGACCTAAATCCACGACTCTGCTTAGCGCTTTGGCG 1830
XX
XX QY 236 gccgcgtcgtgtttaaagcctgattgacctgagagatattg 337
XX || | | | | | | | | | | | | | | | | | | | | | |
XX DB 1829 TTATCGGTGTCTTTAAATGTCGCTTGCTTGGCTTTTTCG 1788
XX
XX RESULT 13
XX ABL23687
XX ID ABL23687 standard; DNA: 2307 BP.
XX
XX AC ABL23687;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 22534.
XX
XX KW Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX MO200171042-A2.
XX
XX PD 27-SEP-2001.
XX

```

```

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 22534; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB57737-AB572072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 2307 BP; 524 A; 665 C; 647 G; 471 T; 0 other;
XX
XX Query Match 3.1%; Score 34.8; DB 23; Length 2307;
XX Best Local Similarity 53.7%; Pred. NO. 5.4;
XX Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
XX QY 401 tgcgtgacctggcggtgagacctaccatcgctcctcctgattctgcactcaacgcgg 460
XX ||||| | | | | | | | | | | | | | | | | | | | | | |
XX DB 611 tgcgtgagcagcaggacaagaccgctccagcagcgttcgctgcggtccatacagcag 670
XX
XX QY 461 ttgaaaaatgcgaactctggtgagtaactggttgatgattgataagattag 520
XX || | | | | | | | | | | | | | | | | | | | | | |
XX DB 671 aggaatcatatgagacacagtgccacgaagggccatcgaaagtgcacactgttcctacg 730
XX
XX QY 521 aaccgaacgcagc 534
XX || | | | | | |
XX DB 731 agctgtacgcaac 744
XX
XX RESULT 14
XX ABL23686/C
XX ID ABL23686 standard; DNA: 5479 BP.
XX
XX AC ABL23686;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 22531.
XX
XX KW Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX MO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX

```



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 11:49:38 ; Search time 72.13 Seconds  
(without alignments)  
3871.971 Million cell updates/sec

Title: US-09-579-383-1  
Perfect score: 1137  
Sequence: 1 atgaatttcacgtaaataa.....tcgtgaagcctgcactaa 1137

Scoring table: IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.4	2.9	2149	1 US-08-784-651-3	Sequence 3, Appli
2	33.4	2.9	7218	1 US-08-222-463-14	Sequence 14, Appl
3	33	2.9	3533	2 US-08-476-062A-40	Sequence 40, Appl
4	33	2.9	3533	5 PCT-US96-01314-40	Sequence 40, Appl
5	33	2.9	5424399-1	Patent No. 5424399	Sequence 7, Appli
6	33	2.9	4080	1 US-08-570-311-7	Sequence 7, Appli
7	33	2.9	4080	2 US-08-353-485-7	Sequence 7, Appli
8	33	2.9	4510	1 US-08-570-311-1	Sequence 1, Appli
9	33	2.9	4510	2 US-08-353-485-1	Sequence 25, Appl
10	33	2.9	6241	1 US-08-570-311-25	Sequence 9, Appli
11	33	2.9	6895	1 US-08-570-311-9	Sequence 9, Appli
12	33	2.9	7266	2 US-08-353-485-9	Sequence 9, Appli
13	33	2.9	7266	3 US-08-336-308A-9	Sequence 9, Appli
14	33	2.9	7266	3 US-08-822-324-5	Sequence 5, Appli
15	33	2.9	7266	4 US-08-450-931-9	Sequence 9, Appli
16	33	2.9	8640	1 US-08-570-311-28	Sequence 28, Appl
17	32.6	2.9	5163	1 US-08-700-651-1	Sequence 4, Appli
18	32.6	2.9	5163	3 US-08-928-361B-4	Sequence 4, Appli
19	32.6	2.9	5318	3 US-08-700-651-2	Sequence 2, Appli
20	32.6	2.9	5318	3 US-08-928-361B-3	Sequence 3, Appli
21	32.4	2.8	2277	1 US-08-676-967-2	Sequence 2, Appli
22	32.4	2.8	2277	1 US-08-676-974-2	Sequence 2, Appli
23	32.4	2.8	2277	2 US-09-098-487-2	Sequence 2, Appli
24	32.2	2.8	2657	2 US-08-974-565C-8	Sequence 8, Appli
25	32.2	2.8	2657	3 US-09-255-748-8	Sequence 8, Appli
26	31.8	2.8	1465	4 US-09-338-671-1	Sequence 1, Appli
27	31.4	2.8	1318	1 US-08-570-311-21	Sequence 21, Appl

28	31.4	2.8	1350	1 US-08-570-311-15	Sequence 15, Appl
29	31.4	2.8	1368	1 US-08-570-311-17	Sequence 17, Appl
30	31.4	2.8	1368	1 US-08-570-311-19	Sequence 19, Appl
31	31.4	2.8	1536	3 US-09-089-195-1	Sequence 1, Appli
32	31.4	2.8	8588	1 US-08-570-311-13	Sequence 13, Appl
33	31.4	2.8	18596	4 US-09-318-448-11	Sequence 11, Appl
34	31.4	2.8	43676	3 US-09-356-952-12	Sequence 12, Appl
35	30.8	2.7	1310	3 US-08-776-900C-22	Sequence 22, Appl
36	30.8	2.7	1310	4 US-09-268-195C-22	Sequence 22, Appl
37	30.8	2.7	1414	1 US-08-224-930-3	Sequence 3, Appli
38	30.8	2.7	3139	1 US-07-894-212A-1	Sequence 1, Appli
39	30.6	2.7	633	4 US-08-998-416-161	Sequence 161, App
40	30.6	2.7	4849	2 US-08-540-804-13	Sequence 13, Appl
41	30.6	2.7	4849	2 US-08-218-265-13	Sequence 13, Appl
42	30.6	2.7	4849	3 US-08-521-872-13	Sequence 13, Appl
43	30.6	2.7	4849	4 US-08-590-399-13	Sequence 13, Appl
44	30.6	2.7	6418	1 US-08-480-528A-11	Sequence 11, Appl
45	30.6	2.7	6418	1 US-08-479-666-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-08-784-651-3  
: Sequence 3, Application US/08784651  
: Patent No. 5821102  
GENERAL INFORMATION:  
: APPLICANT: Berka, Randy  
: APPLICANT: Boomlathnan, Karuppan  
: APPLICANT: Sandal, Thomas  
: TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides  
: NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
: ADDRESSEE: No. 58211020 No. 5821102disk of No. 5821102th America, Inc.  
: STREET: 405 Lexington Avenue  
: CITY: New York  
: STATE: NY  
: COUNTRY: USA  
: ZIP: 10174  
COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/784,651  
: FILING DATE: 21-JAN-1997  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Agtis, Cheryl H.  
: REGISTRATION NUMBER: 34,086  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 212-867-0123  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2149 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: Genomic DNA  
: US-08-784-651-3

Query Match 2.9%; Score 33.4; DB 1; Length 2149;  
Best Local Similarity 54.5%; Pred. No. 0.91;  
Matches 67; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 tccttgatctgcactaacgctgtgaataatcgcaatcctggtgatgaactggc 495  
DB 1877 tcgttctctcaatcaatctggtctatcaatcgttaataatcaatgaagaaacagg 1936

```

QY      496   tctgagtgatgatgttagattacgcgaaccgaaggcgttggatggtctgaattgataaa 555
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1937   TAAATGCAATGCTTTTGTACAAAACCATTGAATGGCTTTTATTATAGATGAAGAATAAC 1996

QY      556   gaa 558
           ||
Db      1997   CAA 1999

RESULT      2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: PALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match          2.9%; Score 33.4; DB 1; Length 7218;
Best Local Similarity 8.8%; Pred.No.2;
Matches 40; Conservative 203; Mismatches 214; Indels 0; Gaps 0.

QY      530   gaaccgaagcgactttagtgtcgaatgataagaanaaacgcgaacttcctttgacaa 579
           ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1392   RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1333

QY      580   taagtgcaccaaatcgcgcaatacatatgttgtatgataaactcagatcacgacatcc 639

```

```

Db      1332 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1273
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      640 tctatgcgctcgtgagctgcatacggttcaacgccgaaaaactgtatgatcac 699
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1272 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1213
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      700 gaagtcgcgtlaaacagcaaatlcttaacaacccggacgttaaagaanaactgttacgc 759
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1212 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1153
           : : : : ~~~~~~ : : : : : : : : : : : : : : : : : :
Qy      760 gcagcccagcatgatctgcgggtgtygtccatatcctgatgaacacctgaaagatacy 819
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1152 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1093
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      820 atgacatgagtgtttgtgcaggatccaactacaagaactaccgatcttcgatttg 879
           : : : : : : : : : : : : : : : | | | | | | | | | |
Db      1092 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTGCGACGCCAACACTGG 1033
           : : : : ~~~~~~ : : : : : : : : : : : : : : : : : :
Qy      880 aaagaattatcagctcctcagcagcttatcatgtgtaaaa 916
           || |||| | || ||| | | | || || |
Db      1032 AATTAA TTCGTG AGCCTA TGGCAA CGACGA GAAAAA 996
                                           |||||

RESULT          3
US-08-476-062A-40/C
Sequence 40 Application US/08476062A
Patent No. 5877275
GENERAL INFORMATION:
APPLICANT: Arnaout, M. Amin
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO.: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
```

NAME/KEY: Coding Sequence  
LOCATION: 75...3530  
US-08-476-062A-40

Query Match 2.9%; Score 33; DB 2; Length 3533;  
Best Local Similarity 58.8%; Pred. No. 1.7;  
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
OY 1024 gaaacaactgaataaacgcgcagatggttcgacctgtgtctctgtccagcgacaac 1083  
DB 3470 GAAGCCGAGCTTGTACACCGCGCGGTGATGAGGCGCAGGACGACGACTCCCGCAGACA 3411  
OY 1084 ggcgcacataatgaacagctgagatcgactctcg 1120  
DB 3410 GCTGCCACGATGAGCGCGGCGGTGCGGACCTCG 3374

RESULT 4  
PCT-US96-01314-40/c  
Sequence 40, Application PC/TUS9601314  
GENERAL INFORMATION:  
APPLICANT: M. Amin Arnaout  
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167  
FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3533 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US96-01314-40

Query Match 2.9%; Score 33; DB 5; Length 3533;  
Best Local Similarity 58.8%; Pred. No. 1.7;  
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
OY 1024 gaaacaactgaataaacgcgcagatggttcgacctgtgtctctgtccagcgacaac 1083  
DB 3470 GAAGCCGAGCTTGTACACCGCGCGGTGATGAGGCGCAGGACGACGACTCCCGCAGACA 3411  
OY 1084 ggcgcacataatgaacagctgagatcgactctcg 1120  
DB 3410 GCTGCCACGATGAGCGCGGCGGTGCGGACCTCG 3374

RESULT 5  
5424399-1/c  
Patent No. 5424399  
APPLICANT: ARNAOUT, M. AMIN  
TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/78,871  
FILING DATE: 16-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 539,842  
FILING DATE: 18-JUN-1990  
APPLICATION NUMBER: 212,573  
FILING DATE: 28-JUN-1988  
SEQ ID NO. 1:  
LENGTH: 3533  
5424399-1

Query Match 2.9%; Score 33; DB 6; Length 3533;  
Best Local Similarity 58.8%; Pred. No. 1.7;  
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
OY 1024 gaaacaactgaataaacgcgcagatggttcgacctgtgtctctgtccagcgacaac 1083  
DB 3470 GAAGCCGAGCTTGTACACCGCGCGGTGATGAGGCGCAGGACGACGACTCCCGCAGACA 3411  
OY 1084 ggcgcacataatgaacagctgagatcgactctcg 1120  
DB 3410 GCTGCCACGATGAGCGCGGCGGTGCGGACCTCG 3374

RESULT 6  
US-08-570-311-7  
Sequence 7, Application US/08570311  
Patent No. 5824791  
GENERAL INFORMATION:  
APPLICANT: Proguiske-Fox, Ann  
APPLICANT: Tumwasorn, Somying  
APPLICANT: Lepine, Guylaine  
APPLICANT: Han, Naimeung  
APPLICANT: Lantz, Marilyn  
APPLICANT: Patil, Joseph  
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ted W. Whitlock  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/570,311  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/353,485  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/647,119  
FILING DATE: 25-JAN-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640  
FILING DATE: 08-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Whitlock, Ted W.  
REGISTRATION NUMBER: 36,965  
REFERENCE/DOCKET NUMBER: UF15.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (904) 375-8100  
TELEFAX: (904) 372-5800  
INFORMATION FOR SEQ. ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4080 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 87..3347  
US-08-570-311-7

Query Match 2.98; Score 33; DB 1; Length 4080;  
Best Local Similarity 48.28; Pred. No. 1.9;  
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 915 aaagtaacgtacgtatcatcatcaggtgtacccctgattgtcccgtaacgagccgttaa 974  
Db 1877 AAAGTACTACTATGACGACGACGAGGTTTCCGGGATGACTATGCGGTGATGATCTC 1936  
Qy 975 ccggaagataaaatgctgtgtaaaatctatcgcgattcgtlaaaacttgaacaaact 1034  
Db 1937 CAAGACGGGACGACGACGAGACTTCACGTTGTTTCAAGAAACGCTTAACGGAAT 1996  
Qy 1035 gaataacgcgcagatggttcgcgcctgtgtctctccagcgacaacgcgcgcataa 1094  
Db 1997 AAATRAAGGGGACGACGAGATTGCTTTCACGGAACCAATGCGCCAAACCTCAAG 2056  
Qy 1095 tgaacagctgagc 1107  
Db 2057 TGTATGATCGAG 2069

RESULT 7  
US-08-353-485-7  
Sequence 7, Application US/08353485  
Patent No. 5830710  
GENERAL INFORMATION:  
APPLICANT: Proguiske-Fox, Ann  
APPLICANT: Tumwasorn, Somying  
APPLICANT: Lepine, Guylaine  
APPLICANT: Han, Naiming  
APPLICANT: Lantz, Marilyn  
APPLICANT: Patti, Joseph  
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes  
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ted W. Whitlock  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,485  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/647,119  
FILING DATE: 25-JAN-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/241,640  
FILING DATE: 08-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Whitlock, Ted W.  
REGISTRATION NUMBER: 36,965  
REFERENCE/DOCKET NUMBER: UF15.C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (904) 375-8100  
TELEFAX: (904) 372-5800  
INFORMATION FOR SEQ. ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4080 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 87..3347  
US-08-353-485-7

Query Match 2.98; Score 33; DB 2; Length 4080;  
Best Local Similarity 48.28; Pred. No. 1.9; Indels 0; Gaps 0;  
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 915 aaagtaacgtacgtatcatcatcaggtgtacccctgattgtcccgtaacgagccgttaa 974  
Db 1877 AAAGTACTACTATGACGACGACGAGGTTTCCGGGATGACTATGCGGTGATGATCTC 1936  
Qy 975 ccggaagataaaatgctgtgtaaaatctatcgcgattcgtlaaaacttgaacaaact 1034  
Db 1937 CAAGACGGGACGACGACGAGACTTCACGTTGTTTCAAGAAACGCTTAACGGAAT 1996  
Qy 1035 gaataacgcgcagatggttcgcgcctgtgtctctccagcgacaacgcgcgcataa 1094  
Db 1997 AAATRAAGGGGACGACGAGATTGCTTTCACGGAACCAATGCGCCAAACCTCAAG 2056  
Qy 1095 tgaacagctgagc 1107  
Db 2057 TGTATGATCGAG 2069

RESULT 8  
US-08-570-311-1  
Sequence 1, Application US/08570311  
Patent No. 5824791  
GENERAL INFORMATION:  
APPLICANT: Proguiske-Fox, Ann  
APPLICANT: Tumwasorn, Somying  
APPLICANT: Lepine, Guylaine  
APPLICANT: Han, Naiming  
APPLICANT: Lantz, Marilyn  
APPLICANT: Patti, Joseph  
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes  
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ted W. Whitlock  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS







Db 4457 CAAGCGGGCAGAAACGGCGGAGACTTCACGGTGTTCAGAAAGCCCTAACGGAAAT 4516  
 Oy 1035 gaataaacgcagatggcttcgcgcctgtgtctctcgcagcagcaacgcgcgcataa 1094  
 Db 4517 AAATAGGGCGGAGCAAGATTGCTCTTCCACGGAAGCAATGGCCCAACCTCAAG 4576  
 Oy 1095 tgaacagctggcg 1107  
 Db 4577 TGTATGATCGAG 4589

RESULT 12  
 US-08-353-485-9  
 ; Sequence 9, Application US/08353485  
 ; Patent No. 5830710  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Progulski-Fox, Ann  
 ; APPLICANT: Tumwasorn, Somying  
 ; APPLICANT: Lepine, Guylaine  
 ; APPLICANT: Han, Nailling  
 ; APPLICANT: Lantz, Marilyn  
 ; APPLICANT: Patti, Joseph  
 ; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes  
 ; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Ted W. Whitlock  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: USA  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/353,485  
 ; FILING DATE: 09-DEC-1994  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/647,119  
 ; FILING DATE: 25-JAN-1991  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/241,640  
 ; FILING DATE: 08-SEP-1988  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Whitlock, Ted W.  
 ; REGISTRATION NUMBER: 36,965  
 ; REFERENCE/DOCKET NUMBER: UP15.C2  
 ; TELEPHONE: (904) 375-8100  
 ; TELEFAX: (904) 372-5800  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6895 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 696..5894  
 ; US-08-353-485-9

Query Match 2.9%; Score 33; DB 2; Length 6895;  
 Best Local Similarity 48.2%; Pred. No. 2.7;  
 Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Oy 915 aaagtaagatcaatgcatcatatggtgctcaacctgatgttccggtccacgcggttaa 974  
 Db 4397 AAAGTACTACTATGACCAACGAGCACTTCCCGGGATCACTATGCGTGATGATCTC 4456  
 Oy 975 ccgcagcagataaagtcgcggaataatcatatcgcgatttcgtaaaactgaaacaact 1034  
 Db 4457 CAAGCGGGCAGAAACGGCGGAGACTTCACGGTGTTCAGAAAGCCCTAACGGAAAT 4516  
 Oy 1035 gaataaacgcagatggcttcgcgcctgtgtctctcgcagcagcaacgcgcgcataa 1094  
 Db 4517 AAATAGGGCGGAGCAAGATTGCTCTTCCACGGAAGCAATGGCCCAACCTCAAG 4576  
 Oy 1095 tgaacagctggcg 1107  
 Db 4577 TGTATGATCGAG 4589

RESULT 13  
 US-08-336-308A-9  
 ; Sequence 9, Application US/08336308A  
 ; Patent No. 6017532  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Travis, James  
 ; APPLICANT: Potempa, Jan S.  
 ; APPLICANT: Barr, Philip J.  
 ; APPLICANT: Pavloff, Nadine  
 ; TITLE OF INVENTION: Porphyromonas gingivalis  
 ; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
 ; STREET: 5370 Manhattan Circle, Suite 201  
 ; CITY: Boulder  
 ; STATE: Colorado  
 ; COUNTRY: US  
 ; ZIP: 80303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/336,308A  
 ; FILING DATE: 08-NOV-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/119,361  
 ; FILING DATE: 10-SEP-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/265,441  
 ; FILING DATE: 24-JUN-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ferber, Donna M.  
 ; REGISTRATION NUMBER: 33,878  
 ; REFERENCE/DOCKET NUMBER: 21-93C  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (303) 499-8080  
 ; TELEFAX: (303) 499-8089  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 7266 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 949..6063  
 ; US-08-336-308A-9

```

; NAME/KEY: CDS
; LOCATION: 949..6063
US-08-822-324-5

```

;  
;  
; ANTI-SENSE: NO

FEATURE: CDS  
NAME/KEY: 949..6063  
LOCATION: 949..6063  
US-09-490-931-9

Query Match 2.98; Score 33; DB 4; Length 7266;  
Best Local Similarity 48.28; Pred. No. 2.7;  
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 915 aaagtaacgattacgtatcatcatcaggtcaccctatgttcccgcccaagccgttcaa 974  
Db 4590 AAAGTACTACTATCATCTACGACGAGGTTTCCGGGATCATATGGGTGATGATCTC 4649  
QY 975 cccgacgataaaatgctgttaaatctatcgcgatttcgtlaaaaactgaaacaaact 1034  
Db 4650 CAAGACGGGGACGACGACGAGACTTCACGGTTGTTTCGAGAGAAAGCCCTAAGGGAAT 4709  
QY 1035 gaataaacgcgcagatgttctggcctgtgtctctgtccagcgacaacgcgcgccataa 1094  
Db 4710 AAATAGGGGGGACGACGAGATTCGGTTTCCACGGAAGCCAAATGCGCCCAAACTCAAG 4769  
QY 1095 tgaacagctggcg 1107  
Db 4770 TGTATGGATCGAG 4782

Search completed: May 7, 2002, 15:09:29  
Job time: 11991 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 11:49:38 ; Search time 1778.63 Seconds  
(without alignments)  
8628.010 Million cell updates/sec

Title: US-09-579-383-1  
Perfect score: 1137  
Sequence: 1 atgaatttaccgttaaata.....tcgtgaagcctgcactaa 1137

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_luv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	39.6	3.5	495	10	B1514134 B160014A
2	39.6	3.5	563	10	B1514040 B160014A
3	39.6	3.5	598	10	B1504891 B170014A
4	38.4	3.4	600	10	B1989791 4051-42 M
5	38.4	3.4	725	10	BG911261 602807526
6	38.4	3.4	772	10	B1110193 602900608
7	38.4	3.4	963	10	BE573275 601333173
8	38	3.3	799	10	BG973651 602844687
9	37.4	3.3	465	10	B1512890 B160010B
10	37.4	3.3	475	10	B1516075 B160021A
11	36.8	3.2	851	12	CNS000A4 AL055901 Drosoph11
12	36.6	3.2	880	10	B1669111 B1669111
13	36.2	3.2	316	12	CNS007VF AL050885 Drosoph11
14	36.2	3.2	401	10	BG126386 EST472032
15	36.2	3.2	407	12	AO902979 GSSC0786
16	36.2	3.2	703	12	AO357632 CTRBt-EI-
17	36.2	3.2	716	10	B1926776 EST546665

18	36.2	3.2	719	10	B1926607
19	36	3.2	313	10	B1072615
20	36	3.2	742	10	B1109582
21	36	3.2	809	12	BH310912
22	36	3.2	898	9	AU176209
23	35.8	3.1	877	12	BH155426
24	35.8	3.1	896	12	BH138875
25	35.8	3.1	923	12	A2543048
26	35.4	3.1	995	12	CNS00731
27	35.2	3.1	625	9	B6333589
28	35.2	3.1	632	10	BG080114
29	35.2	3.1	655	9	BB625765
30	35.2	3.1	679	9	BB642408
31	35.2	3.1	782	9	AU080792
32	35.2	3.1	844	9	A1663310
33	35.2	3.1	933	10	BF975265
34	35.2	3.1	1809	11	AK009954
35	35	3.1	680	9	AU169979
36	35	3.1	1101	12	CNS002EE
37	34.8	3.1	389	9	A1956552
38	34.8	3.1	468	9	AW011984
39	34.8	3.1	481	9	A1930233
40	34.8	3.1	486	9	AW227070
41	34.8	3.1	506	10	W48324
42	34.8	3.1	543	9	AA289281
43	34.8	3.1	565	9	A1036742
44	34.8	3.1	575	10	BF786668
45	34.8	3.1	597	9	AW318758

## ALIGNMENTS

RESULT 1  
LOCUS B1514134 495 bp mRNA linear EST 29-AUG-2001  
DEFINITION B160014A21A12 Bee Brain Normalized Library, B16 Apis mellifera  
cDNA clone B160014A21A12 5', mRNA sequence.  
ACCESSION B1514134  
VERSION B1514134.1 GI:15364508  
KEYWORDS EST.  
SOURCE honeybee  
ORGANISM Apis mellifera  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Apodopterygota; Hymenoptera; Apocrita; Aculeata  
; Apoidea; Apidae; Apis.  
1 (bases 1 to 495)  
Whitefield,C.W., Soares,B., Robertson,H.M., Pardinas,J., Liu,L.,  
Smoller,D. and Robinson,G.E.  
An Expressed Sequence Tag Resource for Studies of Brain and  
Behavior in the Honey Bee  
Unpublished (2001)  
Contact: Gene E. Robinson  
Department of Entomology  
University of Illinois  
505 S. Goodwin Ave., Urbana, IL 61801, USA  
Tel: 217 265 0309  
Fax: 217 244 3499  
Email: generob@life.uiuc.edu

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

This research was funded by the University of Illinois Critical  
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation  
Award in Functional Genomics to G.E. Robinson and an NSF  
Postdoctoral Fellowship in Bioinformatics to C.W. Whitefield.  
PCR Primers  
FORWARD: TAATAGACGACTATAGG  
BACKWARD: ATTACCCCTACTAAG  
Insert Length: 495 Std Error: 0.00  
Plate: B160014A21 row: A column: 12  
Seq primer: AGCGATACATTCACACAGCA  
High quality sequence stop: 495.  
Location/Qualifiers  
1. 495  
/organism="Apis mellifera"

FEATURES  
source

```

/strain="mixed strains of European bees, predominantly
A.m. ligustica"
/db xref="taxon:7460"
/clone="BB160014A21A12"
/clone_lib="Bee Brain Normalized Library, BB16"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab host="DH10B"
/notes="Organ: brain; Vector: pT73-Pac; Site_1: EcoRI;
Site_2: NotI; The BB16 library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806. RNA was
prepared from dissected brains of adult worker bees of
various ages and various behavioral groups."
BASE COUNT      140 a      122 c      111 g      122 t
ORIGIN

```

```

Query Match      3.5%; Score 39.6; DB 10; Length 495;
Best Local Similarity 47.6%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 117; Conservative 0; Mismatches 129;

```

```

OY 858 ctctaccgattctacggttaagaattatagcactctacgcatatgtaaaa 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 CTCGAAGAGATTCTCTGTTTCGTCGACAAATCAGACGACGACGACAAATCTCAAT 84
OY 918 gtaagattatcgatcatatcgttccgcttaccggttccgctccagcgcttaacc 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 ACTCACTTACAAATATCTCAACGACGACGACGACGACGACGACGACGACGACGAC 144
OY 978 gaacgataaattcgtgtaaaatctatcgcggttctgtaaaaactaaacgaagaa 1037
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 ATTGGGACCCCGACGATATATAGCTCTTCGACTTCCAAAGCAGATCAAGACCGCC 204
OY 1038 taaacgcgcagatggttccgctggtctctgtccagcgaacgcggtccataatga 1097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 205 TAATGACGCAATGCGCAATCTGTGCGCTTCGATCCAGCAACAAATTTGCAGCAATGG 264
OY 1098 acagct 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 265 ATGGAT 270

```

```

RESULT 2
BIS14040      563 bp      mRNA      linear      EST 29-AUG-2001
LOCUS
DEFINITION   BB160014A20A12 Bee Brain Normalized Library, BB16 Apis mellifera
CDNA clone BB160014A20A12 5', mRNA sequence.
ACCESSION   BIS14040
VERSION     BIS14040.1 GI:15364414
KEYWORDS    EST.
SOURCE      honeybee.
ORGANISM    Apis mellifera
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata
; Apoidea; Apoidea; Apis.
1 (bases 1 to 563)
Whitfield,C.W., Soares,B., Robertson,H.M., Pardinas,J., Liu,L.,
Smoller,D. and Robinson,G.E.
An Expressed Sequence Tag Resource for Studies of Brain and
Behavior in the Honey Bee
Unpublished (2001)
Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499
Email: gene@life.uiuc.edu
This research was funded by the University of Illinois Critical
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation

```

```

Award in Functional Genomics to G.E. Robinson and an NSF
Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.
PCR primers
FORWARD: TAATGACGCACTATAGG
BACKWARD: ATTACCTCTACTAAG
Insert Length: 563 Std Error: 0.00
Plate: BB160014A20 row: A column: 12
Seq primer: AGCGATACAAATTCACACGGA
High quality sequence stop: 563.
Location/Qualifiers
1. 563

```

```

FEATURES
source
/organism="Apis mellifera"
/strain="mixed strains of European bees, predominantly
A.m. ligustica"
/db xref="taxon:7460"
/clone="BB160014A20A12"
/clone_lib="Bee Brain Normalized Library, BB16"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab host="DH10B"
/notes="Organ: brain; Vector: pT73-Pac; Site_1: EcoRI;
Site_2: NotI; The BB16 library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806. RNA was
prepared from dissected brains of adult worker bees of
various ages and various behavioral groups."
BASE COUNT      158 a      133 c      135 g      135 t
ORIGIN

```

```

Query Match      3.5%; Score 39.6; DB 10; Length 563;
Best Local Similarity 47.6%; Pred. No. 1.9; Indels 129; Indels 0; Gaps 0;
Matches 117; Conservative 0; Mismatches 129;

```

```

OY 858 ctctaccgattctacggttaagaattatagcactctacgcatatgtaaaa 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 CTCGAAGAGATTCTCTGTTTCGTCGACAAATCAGACGACGACGACGACGACGACGAC 84
OY 918 gtaagattatcgatcatatcgttccgcttaccggttccgctccagcgcttaacc 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 ACTCACTTACAAATATCTCAACGACGACGACGACGACGACGACGACGACGACGAC 144
OY 978 gaacgataaattcgtgtaaaatctatcgcggttctgtaaaaactaaacgaagaa 1037
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 ATTGGGACCCCGACGATATATAGCTCTTCGACTTCCAAAGCAGATCAAGACCGCC 204
OY 1038 taaacgcgcagatggttccgctggtctctgtccagcgaacgcggtccataatga 1097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 205 TAATGACGCAATGCGCAATCTGTGCGCTTCGATCCAGCAACAAATTTGCAGCAATGG 264
OY 1098 acagct 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 265 ATGGAT 270

```

```

RESULT 3
BIS04891      598 bp      mRNA      linear      EST 29-AUG-2001
LOCUS
DEFINITION   BB170014A20G05 Bee Brain Normalized/Subtracted Library, BB17 Apis
mellifera CDNA clone BB170014A20G05 5', mRNA sequence.
ACCESSION   BIS04891
VERSION     BIS04891.1 GI:15355265
KEYWORDS    EST.
SOURCE      honeybee.
ORGANISM    Apis mellifera
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata
; Apoidea; Apoidea; Apis.
1 (bases 1 to 598)
Whitfield,C.W., Soares,B., Robertson,H.M., Pardinas,J., Liu,L.,
Smoller,D. and Robinson,G.E.

```



TITLE An Expressed Sequence Tag Resource for Studies of Brain and Behavior in the Honey Bee  
JOURNAL Unpublished (2001)  
COMMENT Contact: Gene E. Robinson  
Department of Entomology  
University of Illinois  
505 S. Goodwin Ave., Urbana, IL 61801, USA  
Tel: 217 244 3499  
Fax: 217 244 3499  
Email: genrob@life.uiuc.edu

This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.  
PCR Primers  
FORWARD: TAATGACGCTCACTAATAGG  
BACKWARD: ATTAACCTCTCACTAAGG

Insert Length: 598 Std Error: 0.00  
Plate: BB170014A20 row: G column: 05  
Seq primer: AGCGATACCAATTTCACACGGA  
High quality sequence stop: 598.

FEATURES  
source Location/Qualifiers  
1..598  
/organism="Apis mellifera"  
/strain="mixed strains of European bees, predominantly A.m. ligustica"  
/db\_xref="taxon:7460"  
/clone="BB170014A20C05"  
/clone\_lib="Bee Brain Normalized/Subtracted Library, BB17"  
/sex="female"  
/tissue\_type="brain"  
/dev\_stage="adult worker honey bee"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pT73-Pac; Site: 1; EcORI; Site 2: NotI; This BB17 cDNA library was generated by subtraction of the BB16 library with 4000 previously sequenced clones. The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

BASE COUNT 173 a 140 c 145 g 140 t  
ORIGIN

Query Match 3.5%; Score 39.6; DB 10; Length 598;  
Best Local Similarity 47.6%; Pred.No. 2;  
Matches 117; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
QY 858 ctctaccgattctaccggtatgaagaattatagcactctaccgatactatggttaaaa 917  
DB 218 CTCGAGAGATTCTCGTTTCGTCGACATACGACACGACGACGACATCCCTCAAT 277  
QY 918 gtaagattagtgatcatcatggtggttaccgtgattgtccggtccggttcaacc 977  
DB 278 ACTGACTTACAAATACCTCAACGCTACAGATGCTGCGATTCATGCTGCCATCC 337  
QY 978 gaagataaattgtgtgaaatcatatctatcggtatgtgtaaaactgaaactctaa 1037  
DB 338 ATTGGGACCCCGACGAATATGAGCTCTTGACTTCCAAAGCAAGATCAAGACCCG 397  
QY 1038 taacgagcagatggttcggtgtgtgtctcttcacgagcaaacgagccataatga 1097  
DB 398 TAAAGACGCGCAATGCAACATTCGTGCCCTTCATCCACGACCAACATTTGCAGCA 457  
QY 1098 acaagt 1103  
DB 458 ATGAT 463

RESULT 4  
BI989791

LOCUS BI989791 600 bp mRNA linear EST 20-DEC-2001  
DEFINITION 4051-42 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,  
ACCESSION BI989791  
VERSION BI989791.1 GI:17960794  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mu,X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W., White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.  
TITLE Gene expression in the developing mouse retina by EST sequencing and microarray analysis  
JOURNAL Nucleic Acids Res. 29 (24), 4983-4993 (2001)  
COMMENT Department of Biochemistry and Molecular Biology  
University of Texas M.D. Anderson Cancer Center  
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA  
Tel: 713 792 3646  
Fax: 713 790 0329

FEATURES  
source Location/Qualifiers  
1..600  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="Mouse E14.5 retina lambda ZAP II Library"  
/tissue\_type="neural retina"  
/dev\_stage="embryonic day 14.5 post-fertilization"

BASE COUNT 133 a 154 c 185 g 128 t  
ORIGIN

Query Match 3.4%; Score 38.4; DB 10; Length 600;  
Best Local Similarity 50.0%; Pred.No. 4.5;  
Matches 96; Conservative 0; Mismatches 96; Indels 0; Gaps 0;  
QY 674 accgaaataaactctgtatgtagtgaagctccggtatatacaaatattcaacaac 733  
DB 275 ACCTCGACGACCCACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 334  
QY 734 cggagcttaagaagaactgttaacgagcagccagatgagcatctgggtgtgcatct 793  
DB 335 AGCCCGAGGAAATGACATCTGACGCGGCTGTTCTGATACCCGCTTGCGCATG 394  
QY 794 acctatgaacacacctgaagaatgatgatgattgtgtgttgagagattcaactaa 853  
DB 395 ATCTCGACGACGATGATCCAGCTGTGTGTGACACGCTGACTGCGAGATGCCATG 454  
QY 854 cgaactctacg 865  
DB 455 GAAAGACCCACAG 466

RESULT 5  
BI9911261  
LOCUS BG911261 725 bp mRNA linear EST 05-JUN-2001  
DEFINITION 602807526P1 NCI-GAP\_Brn67 Homo sapiens cDNA clone IMAGE:4939710  
5', mRNA sequence.  
ACCESSION BG911261  
VERSION BG911261.1 GI:14291737  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.





Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G., and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups.

BASE COUNT

133 a 115 c 101 g 116 t

Query Match 3.3%; Score 37.4; DB 10; Length 465;  
Best Local Similarity 47.3%; Pred. No. 8;  
Matches 113; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

ORIGIN

Qy 865 gattccaccgtatagaagaattatatacgaactcctacgcatatctggttaaaagtagcat 924  
Db 3 GATTCTCGCTTTTCGTCGACATACGACACGACGACGACATCTCAATATCTACT 62  
Qy 925 tacgtatcatcatcattgggttcacccgtatgttcccgctcaccgacctcaaccgaaagat 984  
Db 63 TACAAATACCTCCAAACGGTACAGATGGCTGTGCATTCATGTCTCCATCCATTCGGC 122  
Qy 985 aaatgctgttaaatctatcgcgcgattcgtaaaactaaacaaactgaataaagc 1044  
Db 123 ACCCCAGCAATATGAGCTCTTCTGACTTCCAAAGCAAGATCAAGACCGCTTAATGAC 182  
Qy 1045 gcaagatggttcggcctgtgtctctgtccagcagacagcgccataatgaacagct 1103  
Db 183 GCGAATGCGCAACATCTGTCTGCTGCATCCAGCAGCAATTCGACCAATGATGAT 241

RESULT 10

LOCUS B1516075 475 bp mRNA linear EST 29-AUG-2001  
DEFINITION B160021A20C09 Bee Brain Normalized Library, B16 Apis mellifera  
CDNA clone B160021A20C09 5', mRNA sequence.

ACCESSION B1516075  
VERSION B1516075.1 GI:15366449

KEYWORDS EST.  
SOURCE honeybee.

ORGANISM

Apis mellifera  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata  
; Apoidea; Apidae; Apis.

REFERENCE 1 (bases 1 to 475)

AUTHORS

WHITEFIELD, C.W., SOARES, B., ROBERTSON, H.M., PARDINAS, J., LIU, L.,  
SMOLLER, D., and ROBINSON, G.E.

TITLE

An Expressed Sequence Tag Resource for Studies of Brain and  
Behavior in the Honey Bee

UNPUBLISHED (2001)

CONTACT: Gene E. Robinson

Department of Entomology

University of Illinois

505 S. Goodwin Ave., Urbana, IL 61801, USA

Tel: 217 265 0309

Fax: 217 244 3499

Email: generobi@life.uiuc.edu

This research was funded by the University of Illinois Critical  
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation  
Award in Functional Genomics to G.E. Robinson and an NSF  
Postdoctoral Fellowship in Bioinformatics to C.W. Whitefield.

PCR Primers

FORWARD: TAAATGACATCTACTATAGG

BACKWARD: ATTAACCTCTACTAAG

Insert Length: 475 Std Error: 0.00

Plate: B160021A20 row: C column: 09

Seq primer: AGCGATACATTCACACAGCA

High quality sequence stop: 475.

Location/Qualifiers

1. 475

/organism="Apis mellifera"

/strain="mixed strains of European bees, predominantly  
A.m. ligustica"

/db\_xref="taxon:7460"

/clone="B160021A20C09"  
/clone\_11b="Bee Brain Normalized Library, B16"  
/sex="female"  
/tissue="brain"  
/dev\_stage="adult worker honey bee"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: p7773-Pac; Site: 1; EcorI;  
Site: 2; NotI; The B16 library was constructed by the  
Soares laboratory and it was constructed and normalized  
as described by Bonaldo, M.F., Lennon, G., and Soares,  
M.B. (1996), Genome Research 6(9): 791-806. RNA was  
prepared from dissected brains of adult worker bees of  
various ages and various behavioral groups."

BASE COUNT

136 a 117 c 106 g 116 t

Query Match 3.3%; Score 37.4; DB 10; Length 475;  
Best Local Similarity 47.3%; Pred. No. 8;  
Matches 113; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

ORIGIN

Qy 865 gattccaccgtatagaagaattatatacgaactcctacgcatatctggttaaaagtagcat 924  
Db 12 GATTCTCGCTTTTCGTCGACATACGACACGACGACGACATCTCAATATCTACT 71  
Qy 925 tacgtatcatcatcattgggttcacccgtatgttcccgctcaccgacctcaaccgaaagat 984  
Db 72 TACAAATACCTCCAAACGGTACAGATGGCTGTGCATTCATGTCTCCATCCATTCGGC 131  
Qy 985 aaatgctgttaaatctatcgcgcgattcgtaaaactaaacaaactgaataaagc 1044  
Db 132 ACCCCAGCAATATGAGCTCTTCTGACTTCCAAAGCAAGATCAAGACCGCTTAATGAC 191  
Qy 1045 gcaagatggttcggcctgtgtctctgtccagcagacagcgccataatgaacagct 1103  
Db 192 GCGAATGCGCAACATCTGTCTGCTGCATCCAGCAGCAATTCGACCAATGATGAT 250

RESULT 11

CNS00A0A

LOCUS

DEFINITION

851 bp DNA linear GSS 03-JUN-1999

Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR21C19 of RPCR-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

AL0555901

AL0555901.1 GI:4936670

GSS.

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 851)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammosser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCR-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

p1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES	source	location/Qualifiers	1. 851	/organism="Drosophila melanogaster"	/db_xref="taxon:7227"	/clone_1lb="RPC1-98"	/clone="BACR21C19"	/note="end : TET3"
BASE COUNT	255 a	182 c	167 g	232 t	15 others			
ORIGIN								
Query Match	3.2%	Score 36.8;	DB 12;	Length 851;				
Best Local Similarity	56.0%;	Pred. No. 15;						
Matches	65;	Conservative	2;	Mismatches	49;	Indels	0;	Gaps 0;
OY	877	atgaagaattatcacgactcctcgaactactatggtlaaagaatcagatcagatc	936					
Db	431	ATGAAAACCTTATTAATAACACACACATCAACTTCCTCAAAACACTTGACATTCATGATGTGA	490					
OY	937	atgggcctaccctatgttccgcgcacgcgcgttcacacccgaagaataatgct	992					
Db	491	ATTGGCATTCATCGCAGATACCTTGCTCTAGTACTTAAGCATTTGAATAAGATGCT	546					
RESULT 12								
Bi669111								
LOCUS	603395430P1	NIH_MGC_90	Homo sapiens	CDNA	IMAGE:5405334.5,			
DEFINITION								
ACCESSION	Bi669111							
VERSION	Bi669111.1	GI:16042784						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.							
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cgapbs-remail.nih.gov							
	Tissue Procurement: ATCC							
	CDNA Library Preparation: Life Technologies, Inc.							
	Genomic Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)							
	DNA sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNLN at:							
	http://image.lnln.gov							
	Plate: LLM12033	row: 1	column: 07					
FEATURES								
source								
	High quality sequence stop: 845.							
	location/Qualifiers							
	1. 880							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:5405334"							
	/clone_1lb="NIH_MGC_90"							
	/russue_type="adenocarcinoma, cell line"							
	/lab_host="DH10B (phage-resistant)"							
	/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;							
	/site_2: SalI; Cloned unidirectionally; oligo-dT primed.							
	Average insert size 1.7 kb. Library enriched for							
	full-length clones and constructed by Life Technologies.							
	Note: this is a NIH_MGC Library."							
BASE COUNT	194 a	214 c	238 g	234 t				
ORIGIN								
Query Match	3.2%	Score 36.6;	DB 10;	Length 880;				
Best Local Similarity	46.9%;	Pred. No. 18;						
Matches	114;	Conservative	0;	Mismatches	129;	Indels	0;	Gaps 0;
OY	280	gaatcatcgtaggtagccgctgctgtttaaagcctgattgcttgcgtgagatatatgt	339					

D	b	623	GAAAGCTGCGTCTGATTACCTACAGCTTCAGAGGCCAGTTGCCCCCAAGTCATTTGGG	682
O	y	340	ctgaatgagtaactcaatgatgcataatcgtcgcaaaagctcgtccggacattatly	399
D	b	683	TGGGTGCCTGTCTTTCTTCGTGGATTAATAGCTTGACTGCATCAGTGTGAACAACAGCTGTGA	742
O	y	400	cgcgtgagccttggcggtgtagaacaccatccgtccttccttgatcttcgaatcaagcg	459
D	b	743	AGCTCAGACCTGGAGGTGAGGCTTCCTGAGGCTTTGGCCCCATTATCCACCCCAAAGATTTC	802
O	y	460	gttgaaaaactgcagaatctgtgtgtagaacctcggcttcgagtcatatgtagattac	519
D	b	803	GTTGGCTCGAGCTCAGTTCAGTGTGGTTAAGACACTGGGAGAGAAGAATTCTGCTTAAGTT	862
O	y	520	gaa 522	
D	b	863	GAA 865	
RESULT 13				
C	N	S007VF	CNS007VF	
L	O	CUS	CNS007VF	
D	E	Finition	CNS007VF	316 bp DNA linear GSS 03-JUN-1999
A	C	cession	Drosophila melanogaster genome survey sequence TETJ end of BAC #	
V	E	sion	BACR1612 of RPCL-98 library from Drosophila melanogaster (fruit	
K	E	ywords	fly), genomic survey sequence.	
S	R	ce	AI050885	
O	R	anism	GI:4932939	
A	U	thors	GSS:	
J	I	urnal	fruit fly.	
C	O	ment	Drosophila melanogaster	
			Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
			Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
			Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
			1 (bases 1 to 316)	
			Genoscope.	
			Direct Submission	
			Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :	
			BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr	
			- Web : www.genoscope.cns.fr)	
			Determination of this BAC-end sequence was carried out as part of a	
			collaboration with the Berkeley Drosophila Genome Project (BDGP).	
			The BDGP is constructing a physical map of the Drosophila	
			melanogaster genome using these BACs. For further information	
			please see http://www.fruitfly.org The BDGP Drosophila	
			melanogaster BAC library was prepared by Kazutoyo Osoegawa and	
			Aaron Mamoser in Pieter de Jong's laboratory in the Department of	
			Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,	
			NY. The library is named RPCL-98 and was constructed by partial	
			EcoRI digestion of Drosophila DNA provided by the BDGP from the	
			isogenic strain Y2; cn bw sp, the same strain used for the BDGP's	
			pI and EST libraries. A more detailed description of the library	
			and how to order individual BAC clones, the entire library, or	
			filters for hybridization from the BACPAC Resource Center can be	
			found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
F	E	A	Location/Qualifiers	
S	O	URCE	1..316	
			/organism="Drosophila melanogaster"	
			/db_xref="taxon:7227"	
			/clone_1lb="RPCL-98"	
			/clone="BACR1612"	
			/note="end : TETJ"	
B	A	S	BASE COUNT	
O	R	I	GIN	
			120 a 32 c 40 g 48 t 76 others	
Q	U	E	RY MATCH	
			Best Local Similarity 38.4%; Pred. No. 15;	
			Matches 68; Conservative 26; Mismatches 83; Indels 0; Gaps 0;	
O	Y	865	gattcaccgttatgaaagaattatagacctctcagcatactatggtanaagtagcat	924
D	B	33	GANANNACCAGATRNCGATTCARCTGCAGCAGCACRRMGAGGTACGATAAAMWMTAAACGR	92

```

QY 925 tacgtgatacatgagcttcacccatgttccgcgcacgcgttcacccagcat 984
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 93 GWTGTGATCGAATCGTMTTACCCCTGACCTCATCTCGATGAGTCTTCTRTAAMRCM 152
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 985 aaaaatcgtgataaatatcatcgagcatcttgtaaaaactgaacaaactgaataa 1041
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 TGCWTMTGTAAAMAAWMAAAMCAGCMGTGAGAAAAMAAAMATTAAMWTAAATAA 209
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
LOCUS BG126386 401 bp mRNA linear EST 31-JAN-2001
DEFINITION EST472032 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOF12M9 5' sequence, mRNA sequence.
ACCESSION BG126386
VERSION BG126386
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 401)
AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
Hansen, C., Roming, C. and Tanksley, S.
JOURNAL Generation of ESTs from tomato shoot/meristem tissue
COMMENT Unpublished (2001)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
source
1..401
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOF12M9"
/clone_1lb="tomato shoot/meristem"
/issue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI. Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 123 a 63 c 98 g 117 t
ORIGIN

Query Match 3.2%; Score 36.2; DB 10; Length 401;
Best Local Similarity 55.0%; Pred. No. 17;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 439 ttccgatttcgcacacacgcggtgtaaaaactgcaactgtgtgtagtaacgtgctc 498
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 TGCCTTTTGCACGCAAGATGAAGAAAAAGGGGGAATGTGTGTAGAAATGGAAGAA 268
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 499 gatgtattgatgtagattacgaacccaagcagcgtttgatgtgtcgaatgataaaga 558
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 ATTGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 328
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 559 aaagcgagac 567
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 GAGGTTGAC 337
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
LOCUS AQ090297 407 bp DNA linear GSS 09-JAN-2001
DEFINITION GSS1C07862 Trypanosoma cruzi random genomic library Trypanosoma

```

```

ACCESSION AQ090297
VERSION AQ090297.3 GI:10134812
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
REFERENCE 1 (bases 1 to 407)
AUTHORS Agüero, F., Verdun, R., Frasch, A.C.C. and Sanchez, D.O.
TITLE A random sequencing approach for the analysis of the trypanosoma
JOURNAL cruzi genome: general structure, large gene and repetitive DNA
MEDLINE families, and gene discovery
COMMENT Genome Res. 10 (12), 1996-2005 (2000)
20568489
On Sep 14, 2000 this sequence version replaced gi:9377937.
CONTACT: Sanchez D.O.
Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
San Martin)
Av. Gral Paz entre Albarellos y Constituyentes, INTI edificio 24
CP(1650) San Martin, Prov. de BS AS, Argentina
Tel: 54-11-4580-7255 ext 309
Fax: 54-11-4752-9639
Email: dsanchez@ib.unsam.edu.ar
Sequences were basecalled with phred and vector was masked with
crossmatch (see http://genome.washington.edu). Sequences were then
trimmed from both ends to remove low quality bases and masked
vector.
Seq primer: T7
Class: Shotgun.
FEATURES
source
1..407
Location/Qualifiers
/organism="Trypanosoma cruzi"
/strain="CL-Brener"
/db_xref="taxon:5693"
/clone="G34M20"
/clone_1lb="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
randomly sheared using a nebulizer and the 1 to 2 kb range
was gel purified and cloned into the dephosphorylated
HincII site of the vector"
BASE COUNT 94 a 115 c 102 g 96 t
ORIGIN

Query Match 3.2%; Score 36.2; DB 12; Length 407;
Best Local Similarity 50.9%; Pred. No. 17;
Matches 86; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 56 tgtctactatgttctgtatcgaagtgatcgtgtgtcgggtgtaaccgcaaaa 115
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 339 TGACGGCCCTGCACTCTGTGAGAGAGGTTATGCTACAGGTGGCGGAGTCCCTGTAA 280
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 116 accgcgttgaataatcaaaacgtttaagaatcgcgttgaaggtatcatccagggtact 175
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 AACTCCGACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 220
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 176 atcgtctcgtgtgtgagctataatcaaacactgaaggaactgaaccgaa 224
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 219 AACMAAGCGGGGAACAGCGGTGAGGCCGACAAATCATCATTAACGATAA 171
      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: May 7, 2002, 15:07:41
Job time: 11883 sec

```

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 15:13:52 ; Search time 2185.39 Seconds  
(without alignments)  
10887.518 Million cell updates/sec

Title: US-09-579-383-1

Perfect score: 1137

Sequence: 1 atgaatttaccgtaaata.....tcgtgaagcctgcactaa 1137

Scoring table: OLIGO NWC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- GenEmbl:\*
- 1: gb\_da:\*
  - 2: gb\_htg:\*
  - 3: gb\_in:\*
  - 4: gb\_om:\*
  - 5: gb\_ov:\*
  - 6: gb\_pat:\*
  - 7: gb\_ph:\*
  - 8: gb\_pl:\*
  - 9: gb\_pr:\*
  - 10: gb\_ro:\*
  - 11: gb\_sts:\*
  - 12: gb\_sy:\*
  - 13: gb\_un:\*
  - 14: gb\_vl:\*
  - 15: em\_da:\*
  - 16: em\_fun:\*
  - 17: em\_hum:\*
  - 18: em\_in:\*
  - 19: em\_jnu:\*
  - 20: em\_om:\*
  - 21: em\_or:\*
  - 22: em\_ov:\*
  - 23: em\_pat:\*
  - 24: em\_ph:\*
  - 25: em\_pl:\*
  - 26: em\_ro:\*
  - 27: em\_sts:\*
  - 28: em\_un:\*
  - 29: em\_vl:\*
  - 30: em\_htg\_hum:\*
  - 31: em\_htg\_in:\*
  - 32: em\_htg\_other:\*
  - 33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	ID	Description
------------	-------	--------------	----	-------------

C	73	19	1.7	54525	2	AC103846	AF172445 Plasmidiu
C	72	19	1.7	51919	2	AC100054	AF361355 Rattus no
C	71	19	1.7	40810	8	SPBC1773	AX079087 Sequence
C	70	19	1.7	37412	9	AP000327	AC027117 Homo sapi
C	69	19	1.7	34502	9	AP000327	AC097717 Homo sapi
C	68	19	1.7	29885	9	AC093679	AC015529 Homo sapi
C	67	19	1.7	28045	9	AC012839	AC108032 Homo sapi
C	66	19	1.7	26965	9	AC012839	AC108190 Fells cat
C	65	19	1.7	23049	1	AE008854	AC097081 Rattus no
C	64	19	1.7	20674	1	AE008854	AC108193 Fells cat
C	63	19	1.7	12621	1	AE083352	AC004485 Homo sapi
C	62	19	1.7	4093	1	AF083352	AL596170 Listeria
C	61	19	1.7	4053	1	STY0ADGABA	AF182790 Pueraria
C	60	19	1.7	3693	1	AB015404	AF353629 Trypanoso
C	59	19	1.7	3684	1	AB017921	AF353627 Trypanoso
C	58	19	1.7	3531	1	AB015408	AF353628 Trypanoso
C	57	19	1.7	3531	1	AB015405	AF353626 Trypanoso
C	56	19	1.7	3524	1	AF249275	M34790 T. brucei pr
C	55	19	1.7	3522	1	AF202972	M17027 T. brucei pr
C	54	19	1.7	3519	1	AB015410	L02933 Trypanosoma
C	53	19	1.7	2596	5	DRE315610	X16015 T. brucei pr
C	52	19	1.7	2401	8	AB041754	X52585 T. brucei pr
C	51	19	1.7	1859	8	AB041754	M33130 T. brucei pr
C	50	19	1.7	1800	1	BACCHIDA	X52584 T. brucei pr
C	49	19	1.7	1102	8	AT065408	M33129 T. brucei pr
C	48	19	1.7	1023	8	AY028612	AC099876 Mus muscu
C	47	19	1.7	464	3	AF117752	AC084216 Homo sapi
C	46	19	1.7	405	9	HUMYR94C02	AL356307 Human DNA
C	45	20	1.8	258024	2	AC084046	AC087214 Papio cyn
C	44	20	1.8	258024	2	AC084046	AC079140 Homo sapi
C	43	20	1.8	258024	2	AC084046	AC092199 Trypanoso
C	42	20	1.8	201009	2	AC108475	AC009196 Human Chr
C	41	20	1.8	186927	2	AC078844	AC053466 Homo sapi
C	40	20	1.8	186927	2	AC078844	AC068185 Homo sapi
C	39	20	1.8	176495	2	AC087215	AC087215 Papio cyn
C	38	20	1.8	174189	2	AC068185	AC078844 Homo sapi
C	37	20	1.8	172883	9	AC053466	AC079871 Mus muscu
C	36	20	1.8	170767	9	AC009196	AC108475 Homo sapi
C	35	20	1.8	168706	2	AC092199	AC099285 Rattus no
C	34	20	1.8	161415	9	AC079140	AC084046 Trypanoso
C	33	20	1.8	141415	9	AC087214	AC084046 Trypanoso
C	32	20	1.8	83539	2	AL356307	AF085982 Homo sapi
C	31	20	1.8	62800	2	AC084216	AF117752 Anopheles
C	30	20	1.8	57407	2	AC099876	AY028612 Arabidops
C	29	20	1.8	3073	3	TRBPAPAP	AT065408 Arabidops
C	28	20	1.8	3073	3	TRBPAPAP	DI0594 Bacillus cl
C	27	20	1.8	1712	3	TRBPAPAP	AB041754 Trichoder
C	26	20	1.8	1712	3	TRBPAPAP	M57601 B. circulans
C	25	20	1.8	1712	3	TRBPAPAP	AJ315610 Danio rer
C	24	20	1.8	1267	3	TRBPAPAP	AB015410 Helicobac
C	23	20	1.8	856	3	TRBPAPAP	AF202972 Helicobac
C	22	20	1.8	794	3	TRBPAPAP	AF249275 Helicobac
C	21	20	1.8	763	3	TRBPAPAP	AB015408 Helicobac
C	20	20	1.8	740	3	AF353628	AB015408 Helicobac
C	19	20	1.8	728	3	TRBPAPAP	AB017921 Helicobac
C	18	20	1.8	718	3	TRBPAPAP	AB015404 Helicobac
C	17	20	1.8	718	3	TRBPAPAP	M66434 Salmonella
C	16	20	1.8	636	10	MMU272046	AF083352 Helicobac
C	15	20	1.8	573	8	AF182790	AE008621 Rickettsac
C	14	20	1.8	573	8	AF182790	AE008854 Salmonell
C	13	21	1.8	347050	1	AL596170	AE008731 Salmonell
C	12	21	1.8	313450	1	AL596170	AC012839 Drosophil
C	11	21	1.8	185516	9	AC004485	AC093679 Homo sapi
C	10	21	1.8	139411	2	AC108193	AL023753 Human DNA
C	9	21	1.8	138247	2	AC097081	AP000327 Homo sapi
C	8	21	1.8	133388	2	AC108190	AL03389 S. pombe c
C	7	22	1.9	173998	2	AC108032	AC100054 Mus muscu
C	6	22	1.9	168415	30	AC012529	AC103846 Homo sapi
C	5	22	1.9	142978	9	AC097717	
C	4	22	1.9	69006	2	AC027117	
C	3	22	1.9	2122	10	AF361355	
C	2	22	1.9	630	10	AF361355	
C	1	22	100.0	1137	3	AF172445	

74	19	1.7	64722	2	AC100765	Homo sapi	147	19	1.7	187960	9	AP000866	AP000866 Homo sapi
75	19	1.7	75254	2	AC016176	Homo sapi	148	19	1.7	188176	2	AL513533	AL513533 Homo sapi
76	19	1.7	80429	2	AC022703	Homo sapi	149	19	1.7	189349	8	AC068654	AC068654 Genomic S
77	19	1.7	86022	8	ATAC011623	Arabidops	150	19	1.7	190721	8	AP003263	AP003263 Oryza sat
78	19	1.7	94669	2	AC079113	Homo sapi	151	19	1.7	191824	9	AL161935	AL161935 Human DNA
79	19	1.7	100000	9	AP000055	Homo sapi	152	19	1.7	192370	2	AC016857	AC016857 Homo sapi
80	19	1.7	100000	9	AP000123	Homo sapi	153	19	1.7	199035	2	AC022263	AC022263 Homo sapi
81	19	1.7	100000	9	AP000170	Homo sapi	154	19	1.7	204328	2	AC109929	AC109929 Oryza sat
82	19	1.7	101458	2	AP000680	Homo sapi	155	19	1.7	207309	2	AC099792	AC099792 Homo sapi
83	19	1.7	105198	8	AC001806	Genomic S	156	19	1.7	208393	2	AC087870	AC087870 Mus muscu
84	19	1.7	108718	2	AC008837	Homo sapi	157	19	1.7	210651	2	AC105137	AC105137 Homo sapi
85	19	1.7	110000	2	AL606505_1	Continuation (2 of	158	19	1.7	211348	2	AC099476	AC099476 Rattus no
86	19	1.7	111366	2	AP003817	Oryza sat	159	19	1.7	220044	2	AL603924	AL603924 Mus muscu
87	19	1.7	112782	2	AL359829	Homo sapi	160	19	1.7	221166	3	AE003722	AE003722 Drosophill
88	19	1.7	112926	2	AP004040	Oryza sat	161	19	1.7	245968	2	AL365443	AL365443 Homo sapi
89	19	1.7	115812	9	AC004796	Homo sapi	162	19	1.7	251672	2	AL669951	AL669951 Mus muscu
90	19	1.7	125463	2	AC073881	Homo sapi	163	19	1.7	287571	2	AC092338	AC092338 Homo sapi
91	19	1.7	129206	2	HS022306	Homo sapi	164	19	1.7	287201	2	AC091956	AC091956 Homo sapi
92	19	1.7	129600	2	AL163539	Homo sapi	165	19	1.7	340000	9	AP001720	AP001720 Homo sapi
93	19	1.7	130737	2	AC092360	Homo sapi	166	18	1.6	193	11	G01819	G01819 human STS S
94	19	1.7	132824	2	AC106487	Rattus no	167	18	1.6	230	11	AB019596	AB019596 Homo sapi
95	19	1.7	133641	2	AC106487	Rattus no	168	18	1.6	423	11	G34017	G34017 human STS S
96	19	1.7	136371	9	AC025451	Homo sapi	169	18	1.6	571	9	AB018509	AB018509 Homo sapi
97	19	1.7	137218	2	OSJN00037	AL606617 Oryza sat	170	18	1.6	629	11	G40671	G40671 zebraf
98	19	1.7	138459	2	AP003936	Oryza sat	171	18	1.6	665	8	NTNPL251	NTNPL251 human STS S
99	19	1.7	138714	2	AF377947	Oryza sat	172	18	1.6	951	9	HSFGFR2A19	HSFGFR2A19 Homo sapi
100	19	1.7	140092	2	AP004465	Oryza sat	173	18	1.6	1017	3	AY070925	AY070925 Drosophill
101	19	1.7	141152	2	AP004343	Oryza sat	174	18	1.6	1023	1	AF410852	AF410852 Streptococ
102	19	1.7	144557	2	AP004631	Oryza sat	175	18	1.6	1113	1	SESDC	X17241 S. equistimil
103	19	1.7	144782	2	AC105036	Oryza sat	176	18	1.6	1128	8	AF181939	AF181939 Amphicarp
104	19	1.7	146713	2	AL355795	Human DNA	177	18	1.6	1434	1	AB032593	AB032593 Flexibact
105	19	1.7	147505	2	AP004699	Oryza sat	178	18	1.6	1484	6	AB083516	AB083516 Sequence
106	19	1.7	147992	2	AP004260	Oryza sat	179	18	1.6	1892	1	AY051581	AY051581 Drosophill
107	19	1.7	148697	2	AC104322	Homo sapi	180	18	1.6	2147	1	SPSDACENE	X84793 S. pyogenes
108	19	1.7	148831	2	AC068526	Homo sapi	181	18	1.6	2247	4	AF457134	AF457134 Ovis arlie
109	19	1.7	155341	2	AC104784	Homo sapi	182	18	1.6	2446	4	OC037769	X37769 Oryctolagus
110	19	1.7	156615	9	AL138720	Human DNA	183	18	1.6	2462	9	AK026806	AK026806 Homo sapi
111	19	1.7	156870	2	AC091916	Homo sapi	184	18	1.6	3248	9	HSFGFR2MR	Z71929 H. sapiens F
112	19	1.7	157402	9	AC092333	Homo sapi	185	18	1.6	3248	9	HSFGFR2UA	Z69641 H. sapiens F
113	19	1.7	157450	8	AC079037	Oryza sat	186	18	1.6	3248	9	HSFGFR2UB	Z69640 H. sapiens F
114	19	1.7	158039	2	AC021840	Homo sapi	187	18	1.6	3415	9	HSFGFRBE	X52832 Human bek m
115	19	1.7	160791	2	AC094557	Rattus no	188	18	1.6	4268	6	AX332811	AX332811 Sequence
116	19	1.7	161475	9	HS164C20	Homo sapi	189	18	1.6	4268	6	AX334117	AX334117 Sequence
117	19	1.7	161720	8	AP003281	Oryza sat	190	18	1.6	4268	6	AX336422	AX336422 Sequence
118	19	1.7	163028	9	AC005877	c1tb_19_o	191	18	1.6	4268	6	H0MKSAMI	M87770 Human fibro
119	19	1.7	164502	9	AC012613	Homo sapi	192	18	1.6	4268	6	AX155251	AX155251 Sequence
120	19	1.7	165105	2	AC024240	Homo sapi	193	18	1.6	6614	6	ARI30887	ARI30887 Sequence
121	19	1.7	165490	3	AC009736	Drosophill	194	18	1.6	9291	1	AE009645	AE009645 Brucella
122	19	1.7	166385	2	AL358783	Homo sapi	195	18	1.6	14806	2	AC017285	AC017285 Drosophill
123	19	1.7	166495	2	OSJN00072	AL606639 Oryza sat	196	18	1.6	14867	9	AF102137	AF102137 Homo sapi
124	19	1.7	167399	8	AP003236	AL156843 Homo sapi	197	18	1.6	15852	1	AE007052	AE007052 Mycobacte
125	19	1.7	168149	2	AL156843	Homo sapi	198	18	1.6	16446	8	CREI32478	AJ132478 Chlamydom
126	19	1.7	169418	2	AC026628	Homo sapi	199	18	1.6	16682	9	AB030817	AB030817 Homo sapi
127	19	1.7	170162	9	AC096915	Homo sapi	200	18	1.6	21444	2	AC092149	AC092149 Homo sapi
128	19	1.7	170781	9	AC013473	Homo sapi	201	18	1.6	27302	2	AC020017	AC020017 Drosophill
129	19	1.7	172004	9	AC008482	Homo sapi	202	18	1.6	29333	2	AC025951	AC025951 Staphyloc
130	19	1.7	172561	2	AL603890	Homo sapi	203	18	1.6	30327	2	AC095395	AC095395 Rattus no
131	19	1.7	173493	9	AC093115	Homo sapi	204	18	1.6	33100	1	MTCY9F9	M84498 Mycobacteri
132	19	1.7	174907	2	AC025455	Homo sapi	205	18	1.6	33855	2	AC014874	AC014874 Drosophill
133	19	1.7	176083	2	AC103468	Homo sapi	206	18	1.6	38342	6	AX251503	AX251503 Sequence
134	19	1.7	176701	2	AC107211	Rattus no	207	18	1.6	38342	6	AX344502	AX344502 Sequence
135	19	1.7	177464	2	AP003768	AP003768 Oryza sat	208	18	1.6	39573	3	CEC4134	AF101319 Caenorhabd
136	19	1.7	177822	2	AC058796	Homo sapi	209	18	1.6	39750	3	CEC4134	Z48045 Caenorhabd
137	19	1.7	177985	2	AC027271	Homo sapi	210	18	1.6	44310	2	AC098421	AC098421 Rattus no
138	19	1.7	178480	2	AC022323	Homo sapi	211	18	1.6	49105	2	AC067787	AC067787 Homo sapi
139	19	1.7	178845	2	OSJN00217	AL665012 Oryza sat	212	18	1.6	53067	9	AL356535	AL356535 Human DNA
140	19	1.7	179019	2	AC087297	Homo sapi	213	18	1.6	53855	2	AC105281	AC105281 Homo sapi
141	19	1.7	180277	2	AC104270	Oryza sat	214	18	1.6	54611	2	AC101727	AC101727 Mus muscu
142	19	1.7	181573	2	AC020959	Mus muscu	215	18	1.6	59130	9	AL135904	AL135904 Human DNA
143	19	1.7	183191	2	AC108224	Oryza sat	216	18	1.6	59784	2	AC109016	AC109016 Rattus no
144	19	1.7	183385	2	AC073498	Homo sapi	217	18	1.6	60972	2	AC101241	AC101241 Mus muscu
145	19	1.7	184329	2	AP001896	Homo sapi	218	18	1.6	61405	2	AP003576	AP003576 Oryza sat
146	19	1.7	184447	9	AC012519	Homo sapi	219	18	1.6	61423	10	AL365329	AL365329 Mouse DNA

C 220	18	1.6	61890	2	AC098324	Rattus no	293	18	1.6	137007	2	AP004265	Oryza sat
C 221	18	1.6	63017	9	AL157710	Human DNA	294	18	1.6	138846	2	AC068478	Hom sapi
C 222	18	1.6	63341	2	AL157710	Human DNA	295	18	1.6	139788	2	AC024118	Hom sapi
C 223	18	1.6	65258	2	AC104862	Mus muscu	296	18	1.6	140150	9	AC072057	Hom sapi
C 224	18	1.6	65393	2	AC104577	Hom sapi	297	18	1.6	140770	2	AF252832	Hom sapi
C 225	18	1.6	65377	2	AC104577	Hom sapi	298	18	1.6	140915	2	AC005587	Hom sapi
C 226	18	1.6	65839	8	AC005916	Arabidops	299	18	1.6	143559	2	AL1359083	Hom sapi
C 227	18	1.6	66902	8	AC107799	Mus muscu	300	18	1.6	143675	2	AL161625	Hom sapi
C 228	18	1.6	67326	2	AC073034	Hom sapi	301	18	1.6	146017	2	AP003575	Oryza sat
C 229	18	1.6	67334	2	AC102807	Mus muscu	302	18	1.6	146361	2	AC106669	Rattus no
C 230	18	1.6	67557	2	AC055802	Hom sapi	303	18	1.6	147367	2	AC073966	Hom sapi
C 231	18	1.6	68133	2	AC100012	Mus muscu	304	18	1.6	147534	2	AC068620	Hom sapi
C 232	18	1.6	68133	2	AC100012	Mus muscu	305	18	1.6	149312	2	AC073328	Hom sapi
C 233	18	1.6	70673	8	AC091258	Mus muscu	306	18	1.6	150290	2	AC068428	Hom sapi
C 234	18	1.6	71178	8	AP000377	Arabidops	307	18	1.6	150635	2	AL1365185	Human DNA
C 235	18	1.6	72402	2	AC087760	Hom sapi	308	18	1.6	150676	2	AC107235	Oryza sat
C 236	18	1.6	72402	2	AC087760	Hom sapi	309	18	1.6	150942	9	AC017070	Hom sapi
C 237	18	1.6	74119	2	AC036177	Hom sapi	310	18	1.6	151500	3	AC008355	Hom sapi
C 238	18	1.6	74282	8	AB025506	Arabidops	311	18	1.6	151915	2	AC095033	Hom sapi
C 239	18	1.6	77113	2	AC022331	Hom sapi	312	18	1.6	153304	2	AC092155	Hom sapi
C 240	18	1.6	80569	2	AC105791	Rattus no	313	18	1.6	153776	2	AP003511	Oryza sat
C 241	18	1.6	81298	8	AC004705	Arabidops	314	18	1.6	153825	2	AC016358	Hom sapi
C 242	18	1.6	83827	9	AL157369	Human DNA	315	18	1.6	153936	9	AP001207	Hom sapi
C 243	18	1.6	84157	9	AC003006	Human DNA	316	18	1.6	154354	9	AC006197	Hom sapi
C 244	18	1.6	84157	9	AB028615	Arabidops	317	18	1.6	154620	2	AC105185	Hom sapi
C 245	18	1.6	86380	10	MMU131017	Mus muscu	318	18	1.6	155209	2	AC097796	Rattus no
C 246	18	1.6	87401	2	AC021898	Hom sapi	319	18	1.6	155765	2	AC099636	Hom sapi
C 247	18	1.6	89848	8	AB062086	Arabidops	320	18	1.6	156420	9	AC007510	Hom sapi
C 248	18	1.6	90132	8	AC084165	Arabidops	321	18	1.6	156618	9	AC012362	Hom sapi
C 249	18	1.6	90132	8	AC084165	Arabidops	322	18	1.6	156939	9	AC011233	Hom sapi
C 250	18	1.6	91040	2	AC017292	Hom sapi	323	18	1.6	157147	2	AC106431	Rattus no
C 251	18	1.6	94605	2	AF41048052	Drosophill	324	18	1.6	157198	2	AC024082	Human Chr
C 252	18	1.6	95438	8	AC102993	Hom sapi	325	18	1.6	157382	2	AC009279	Hom sapi
C 253	18	1.6	95854	8	AP003077	Oryza sat	326	18	1.6	157610	9	AC087235	Hom sapi
C 254	18	1.6	96280	9	AC006968	Hom sapi	327	18	1.6	158198	2	AC012013	Hom sapi
C 255	18	1.6	97097	9	AL353740	Human DNA	328	18	1.6	158561	2	AC107058	Hom sapi
C 256	18	1.6	98527	8	AC004044	Arabidops	329	18	1.6	159191	9	AL161430	Human DNA
C 257	18	1.6	99141	2	AC097190	Rattus no	330	18	1.6	159560	2	AC025942	Hom sapi
C 258	18	1.6	99479	2	AL136089	Human DNA	331	18	1.6	159582	2	AC015734	Hom sapi
C 259	18	1.6	100000	2	AP000082	Hom sapi	332	18	1.6	159622	2	AC107876	Hom sapi
C 260	18	1.6	101461	9	AC010658	Drosophill	333	18	1.6	159770	2	AC026654	Hom sapi
C 261	18	1.6	102703	8	AC004910	Hom sapi	334	18	1.6	160145	2	AL161612	Human DNA
C 262	18	1.6	103568	8	CNS078GJ	DNA centr	335	18	1.6	160544	2	HS253907	Hom sapi
C 263	18	1.6	105126	2	AC103421	Rattus no	336	18	1.6	161733	2	AC025797	Hom sapi
C 264	18	1.6	105247	8	AC004969	Hom sapi	337	18	1.6	161799	2	AC002091	Genomic s
C 265	18	1.6	105420	8	AC006268	Arabidops	338	18	1.6	162005	2	AL1356002	Hom sapi
C 266	18	1.6	106753	8	AC012375	Genomic s	339	18	1.6	162390	2	AC076959	Hom sapi
C 267	18	1.6	107106	2	AC096280	Rattus no	340	18	1.6	163326	2	AL161445	Human DNA
C 268	18	1.6	109786	8	FXK24	Arabidops	341	18	1.6	164077	2	AC024398	Hom sapi
C 269	18	1.6	109821	9	HS10C16	Human DNA	342	18	1.6	166008	2	AC096264	Hom sapi
C 270	18	1.6	110000	10	AF162137_1	Continuation (2 of	343	18	1.6	166267	2	AC025942	Hom sapi
C 271	18	1.6	110714	8	FP7P1	AP004321 Oryza sat	344	18	1.6	166471	2	AC091052	Hom sapi
C 272	18	1.6	113349	2	AP007359	Arabidops	345	18	1.6	166543	3	AC005813	Drosophill
C 273	18	1.6	116978	2	AP003915	Hom sapi	346	18	1.6	166545	2	AC097192	Rattus no
C 274	18	1.6	116978	2	AP003915	Oryza sat	347	18	1.6	167098	2	AC023952	Hom sapi
C 275	18	1.6	117694	2	AC084050	Hom sapi	348	18	1.6	167358	2	AC010673	Hom sapi
C 276	18	1.6	118313	2	AC003991	Human BAC	349	18	1.6	167862	2	AC094099	Hom sapi
C 277	18	1.6	119977	2	AC0094827	Rattus no	350	18	1.6	168030	2	AC091462	Hom sapi
C 278	18	1.6	125566	2	AC069447	Mus muscu	351	18	1.6	168589	2	AC095160	Hom sapi
C 279	18	1.6	126686	2	AC093565	Hom sapi	352	18	1.6	169562	2	AC095160	Rattus no
C 280	18	1.6	126962	2	AC109595	Oryza sat	353	18	1.6	169677	2	AL136980	Human DNA
C 281	18	1.6	128309	9	AC025450	Hom sapi	354	18	1.6	172584	2	AC036215	Hom sapi
C 282	18	1.6	128480	9	AL359314	Human DNA	355	18	1.6	170000	2	AC004394	Hom sapi
C 283	18	1.6	130030	9	AC004889	Hom sapi	356	18	1.6	170305	2	AC007465	Hom sapi
C 284	18	1.6	130030	9	AC004889	Hom sapi	357	18	1.6	170326	9	AL158838	Human DNA
C 285	18	1.6	133292	2	OSJN00152	Hom sapi	358	18	1.6	172298	2	AC009556	Hom sapi
C 286	18	1.6	133939	2	AC095273	Oryza sat	359	18	1.6	172584	2	AC098859	Hom sapi
C 287	18	1.6	133939	2	AC095273	Rattus no	360	18	1.6	173435	2	AC015715	Hom sapi
C 288	18	1.6	134716	2	AP0034239	Hom sapi	361	18	1.6	173436	2	AC009988	Hom sapi
C 289	18	1.6	134764	2	AP003305	Hom sapi	362	18	1.6	174939	2	AC0096167	Hom sapi
C 290	18	1.6	135586	2	AC095306	Rattus no	363	18	1.6	175286	2	AC046197	Hom sapi
C 291	18	1.6	136129	2	AC011050	Hom sapi	364	18	1.6	176092	2	AC073252	Hom sapi
C 292	18	1.6	136294	2	AC092781	Oryza sat	365	18	1.6	176167	2	AC109995	Hom sapi

C 366	18	1.6 177568	9	HS388M5	297055 Human DNA S	C 439	18	1.6 239116	9	AP000426	AB0000426 Homo sapi
C 367	18	1.6 177596	2	AC091945	AC091945 Homo sapi	C 440	18	1.6 243499	2	AC016186	AC016186 Homo sapi
C 368	18	1.6 177748	2	AC016185	AC016185 Homo sapi	C 441	18	1.6 245968	2	AL365443	AL365443 Homo sapi
C 369	18	1.6 177791	2	AC104162	AC104162 Homo sapi	C 442	18	1.6 249286	2	AL626776	AL626776 Mus muscu
C 370	18	1.6 177922	2	AC022533	AC022533 Homo sapi	C 443	18	1.6 252925	2	AC021048	AC021048 Homo sapi
C 371	18	1.6 177941	9	AL445423	AL445423 Human DNA	C 444	18	1.6 256242	3	AE003756	AE003756 Drosophill
C 372	18	1.6 178328	9	AC092547	AC092547 Homo sapi	C 445	18	1.6 277603	2	AC079543	AC079543 Mus muscu
C 373	18	1.6 178434	2	AC097302	AC097302 Rattus no	C 446	18	1.6 303648	3	AE003486	AE003486 Drosophill
C 374	18	1.6 178642	2	AC023122	AC023122 Homo sapi	C 447	17	1.5 305 11	G05397	G05397 human STS W	
C 375	18	1.6 179086	2	AC053541	AC053541 Homo sapi	C 448	17	1.5 310 6	AX331205	AX331205 Sequence	
C 376	18	1.6 179474	2	AC068898	AC068898 Homo sapi	C 449	17	1.5 321 1	HSANNEYV08	HSANNEYV08	
C 377	18	1.6 179639	2	AC009989	AC009989 Homo sapi	C 450	17	1.5 330 3	AF067721	AF067721 Human annex	
C 378	18	1.6 179642	2	AC099288	AC099288 Rattus no	C 451	17	1.5 333 3	AE067714	AE067714 Heperagr	
C 379	18	1.6 179667	2	AC106895	AC106895 Homo sapi	C 452	17	1.5 344 4	AF093134	AF093134 Thalarcto	
C 380	18	1.6 179819	30	AC024702	AC024702 Homo sapi	C 453	17	1.5 353 11	G27902	G27902 human STS S	
C 381	18	1.6 180354	2	AC009689	AC009689 Homo sapi	C 454	17	1.5 358 4	D55637	D55637 Bos taurus	
C 382	18	1.6 181311	2	AC106862	AC106862 Homo sapi	C 455	17	1.5 370 11	HS703010T	HS703010 H.sapiens	
C 383	18	1.6 181563	2	AL670035	AL670035 Mus muscu	C 456	17	1.5 390 9	U19965	U19965 Homo sapien	
C 384	18	1.6 181660	2	AC087738	AC087738 Homo sapi	C 457	17	1.5 396 3	AF039280	AF039280 Plasmodiu	
C 385	18	1.6 181950	3	AC008093	AC008093 Drosophill	C 458	17	1.5 427 11	G13961	G13961 human STS S	
C 386	18	1.6 181950	3	AC008093	AC008093 Drosophill	C 459	17	1.5 446 1	AF303244	AF303244 Streptoco	
C 387	18	1.6 182147	2	AC013670	AC013670 Homo sapi	C 460	17	1.5 499 5	AF137180	AF137180 Mugil cep	
C 388	18	1.6 182572	2	AC022738	AC022738 Homo sapi	C 461	17	1.5 511 11	G62994	G62994 SHGC-140757	
C 389	18	1.6 182726	2	AC084436	AC084436 Homo sapi	C 462	17	1.5 538 11	G62599	G62599 SHGC-140099	
C 390	18	1.6 183321	2	AC093639	AC093639 Homo sapi	C 463	17	1.5 684 1	SSU64095	SSU64095 Shewanella	
C 391	18	1.6 183370	3	AC008213	AC008213 Drosophill	C 464	17	1.5 688 8	AF349710	AF349710 Russula s	
C 392	18	1.6 183499	9	AC069559	AC069559 Mus muscu	C 465	17	1.5 696 1	NME391278	NME391278 Neisseria	
C 393	18	1.6 183972	9	AC093876	AC093876 Homo sapi	C 466	17	1.5 696 6	NME391280	NME391280 Neisseria	
C 394	18	1.6 184285	9	AC007365	AC007365 Homo sapi	C 467	17	1.5 696 1	NME391280	NME391280 Neisseria	
C 395	18	1.6 186236	9	AC090015	AC090015 Homo sapi	C 468	17	1.5 696 6	AX076276	AX076276 Sequence	
C 396	18	1.6 186292	2	AC064876	AC064876 Homo sapi	C 469	17	1.5 696 6	AX076278	AX076278 Sequence	
C 397	18	1.6 186471	9	AC005053	AC005053 Homo sapi	C 470	17	1.5 696 6	AX076280	AX076280 Sequence	
C 398	18	1.6 187728	2	AC062011	AC062011 Homo sapi	C 471	17	1.5 696 6	AX076282	AX076282 Sequence	
C 399	18	1.6 188170	2	AC107294	AC107294 Homo sapi	C 472	17	1.5 696 6	AX076284	AX076284 Sequence	
C 400	18	1.6 188255	2	AC074241	AC074241 Homo sapi	C 473	17	1.5 696 6	AX076286	AX076286 Sequence	
C 401	18	1.6 188388	2	AC096875	AC096875 Pan trogl	C 474	17	1.5 696 6	AX076288	AX076288 Sequence	
C 402	18	1.6 189271	9	AL138752	AL138752 Human DNA	C 475	17	1.5 696 6	AX076290	AX076290 Sequence	
C 403	18	1.6 189281	2	AC012458	AC012458 Homo sapi	C 476	17	1.5 696 6	AX076292	AX076292 Sequence	
C 404	18	1.6 189335	2	AC087477	AC087477 Homo sapi	C 477	17	1.5 696 6	AX076294	AX076294 Sequence	
C 405	18	1.6 189387	9	AL592287	AL592287 Human DNA	C 478	17	1.5 696 6	AX076296	AX076296 Sequence	
C 406	18	1.6 189572	9	AC091199	AC091199 Homo sapi	C 479	17	1.5 696 6	AX076298	AX076298 Sequence	
C 407	18	1.6 190000	2	AC004580	AC004580 Homo sapi	C 480	17	1.5 696 6	AX076300	AX076300 Sequence	
C 408	18	1.6 190119	2	AC092049	AC092049 Homo sapi	C 481	17	1.5 696 6	AX139587	AX139587 Sequence	
C 409	18	1.6 190215	2	AC016975	AC016975 Homo sapi	C 482	17	1.5 696 6	AX139589	AX139589 Sequence	
C 410	18	1.6 190357	2	AC012259	AC012259 Homo sapi	C 483	17	1.5 696 6	AX139591	AX139591 Sequence	
C 411	18	1.6 190640	2	AP001177	AP001177 Homo sapi	C 484	17	1.5 696 6	AX139593	AX139593 Sequence	
C 412	18	1.6 191281	2	AC013624	AC013624 Homo sapi	C 485	17	1.5 696 6	AX139595	AX139595 Sequence	
C 413	18	1.6 192001	2	AC019139	AC019139 Homo sapi	C 486	17	1.5 696 6	AX139597	AX139597 Sequence	
C 414	18	1.6 192213	2	AC107083	AC107083 Homo sapi	C 487	17	1.5 696 6	AX139599	AX139599 Sequence	
C 415	18	1.6 192530	2	AC090772	AC090772 Homo sapi	C 488	17	1.5 696 6	AX139601	AX139601 Sequence	
C 416	18	1.6 193942	2	AC023983	AC023983 Homo sapi	C 489	17	1.5 696 6	AX139603	AX139603 Sequence	
C 417	18	1.6 195838	2	AC069197	AC069197 Homo sapi	C 490	17	1.5 696 6	AX139605	AX139605 Sequence	
C 418	18	1.6 196517	8	ATCHRIV20	AL161508 Arabidops	C 491	17	1.5 696 6	AX139607	AX139607 Sequence	
C 419	18	1.6 196922	2	AC098158	AC098158 Rattus no	C 492	17	1.5 696 6	AX139609	AX139609 Sequence	
C 420	18	1.6 197788	2	AC012514	AC012514 Homo sapi	C 493	17	1.5 696 6	AX139611	AX139611 Sequence	
C 421	18	1.6 199280	2	AC079364	AC079364 Mus muscu	C 494	17	1.5 771 8	AF343075	AF343075 Neurospor	
C 422	18	1.6 199514	8	ATCHRIV7	AL161495 Arabidops	C 495	17	1.5 783 11	CNS06FC3	AL396217 T3 end of	
C 423	18	1.6 199615	9	ATCHRIV7	AL161495 Arabidops	C 496	17	1.5 821 10	MUS0300183	MUS0300183	
C 424	18	1.6 200123	9	AC073264	AC073264 Homo sapi	C 497	17	1.5 825 6	E03178	E03178 DNA sequenc	
C 425	18	1.6 206497	10	AC084389	AC084389 Mus muscu	C 498	17	1.5 889 9	HUMHY4ROA	HUMHY4ROA	
C 426	18	1.6 209350	2	AC106531	AC106531 Rattus no	C 499	17	1.5 957 6	E05642	E05642 H.sapiens (	
C 427	18	1.6 210515	9	AC097382	AC097382 Homo sapi	C 500	17	1.5 969 6	E36781	E36781 Novel phoH	
C 428	18	1.6 212206	9	AC098482	AC098482 Homo sapi	C 501	17	1.5 978 1	AVCYC553	AVCYC553	
C 429	18	1.6 216841	9	AC025164	AC025164 Homo sapi	C 502	17	1.5 1022 6	AX094133	AX094133	
C 430	18	1.6 217412	2	AC024400	AC024400 Homo sapi	C 503	17	1.5 1041 3	AY071274	AY071274 Drosophill	
C 431	18	1.6 218331	2	AC092180	AC092180 Homo sapi	C 504	17	1.5 1043 8	AF320980	AF320980 aspergill	
C 432	18	1.6 219436	9	AC005046	AC005046 Homo sapi	C 505	17	1.5 1069 8	AF410300	AF410300 Arabidops	
C 433	18	1.6 222885	2	AL158816	AL158816 Homo sapi	C 506	17	1.5 1138 1	AEC2COP5R	AEC2COP5R	
C 434	18	1.6 223793	2	AC023942	AC023942 Homo sapi	C 507	17	1.5 1186 6	AX348476	AX348476 Sequence	
C 435	18	1.6 226534	2	AC099647	AC099647 Homo sapi	C 508	17	1.5 1269 10	AB047002	AB047002	
C 436	18	1.6 227894	3	AE003677	AE003677 Drosophill	C 509	17	1.5 1325 3	AF072442	AF072442 plasmodu-	
C 437	18	1.6 230534	2	AC079575	AC079575 Mus muscu	C 510	17	1.5 1422 1	AB034212	AB034212 Marine ba	
C 438	18	1.6 234411	10	MMAJ2971	AJ297131 Mus muscu	C 511	17	1.5 1422 1	AB047173	AB047173 Marine CF	

512	17	1.5	1422	1	AB047178	AB047178 Marine CF	c 585	17	1.5	7076	1	AFACZCCBAD	D67044 Alcailligenes
c 513	17	1.5	1462	10	AF047043	AF047043 Mus muscu	c 586	17	1.5	7387	8	AF112473	AF112473 Risticella
c 514	17	1.5	1463	10	AF047389	AF047389 Mus muscu	c 587	17	1.5	8620	1	AF065891	AF065891 Risticella
c 515	17	1.5	1585	9	BC000038	BC000038 Homo sapi	c 588	17	1.5	9015	2	AC017826	AC017826 Drosophila
c 516	17	1.5	1600	1	HME239061	AJ239061 Halomonas	c 589	17	1.5	9056	2	CERT2387	U50192 Caenorhabdi
c 517	17	1.5	1627	10	MMSPPHNGO	AJ2222800 Mus muscu	c 590	17	1.5	9276	2	AC018186	AC018186 Drosophila
c 518	17	1.5	1631	10	BC010978	BC010978 Mus muscu	c 591	17	1.5	9650	14	AB024711	AB024711 Gallid he
c 519	17	1.5	1662	10	HSSPHINGO	AJ222801 Homo sapi	c 592	17	1.5	9830	1	U67476	U67476 Methanococc
c 520	17	1.5	1687	1	AF314183	AF314183 Agrobacte	c 593	17	1.5	9942	1	AE009389	AE009389 Agrobacte
c 521	17	1.5	1727	10	AF075266	AF075266 Mus muscu	c 594	17	1.5	9984	1	AE005208	AE005208 Escherich
c 522	17	1.5	1841	10	AF075265	AF075265 Mus muscu	c 595	17	1.5	10006	1	AE002386	AE002386 Neisseria
c 523	17	1.5	1887	3	AY061100	AY061100 Drosophila	c 596	17	1.5	10029	1	AE007775	AE007775 Clostridi
c 524	17	1.5	1890	10	AF075264	AF075264 Mus muscu	c 597	17	1.5	10029	1	AE008226	AE008226 Agrobacte
c 525	17	1.5	1893	10	AF300293	AF300293 Mus muscu	c 598	17	1.5	10111	3	CELC0588	U11029 Caenorhabdi
c 526	17	1.5	1904	10	BC010748	BC010748 Mus muscu	c 599	17	1.5	10171	1	AE007829	AE007829 Clostridi
c 527	17	1.5	1935	8	AF050098	AF050098 Trichoder	c 600	17	1.5	10213	1	AE009763	AE009763 Pyrobact
c 528	17	1.5	1943	10	AF075267	AF075267 Mus muscu	c 601	17	1.5	10310	1	AE007464	AE007464 Streptoc
c 529	17	1.5	1974	10	AF075263	AF075263 Mus muscu	c 602	17	1.5	10374	1	AE009594	AE009594 Brucella
c 530	17	1.5	2004	10	AF075262	AF075262 Mus muscu	c 603	17	1.5	10488	1	AE000894	AE000894 Methanoba
c 531	17	1.5	2009	9	AK057650	AK057650 Homo sapi	c 604	17	1.5	10679	1	AE004575	AE004575 Pseudomon
c 532	17	1.5	2076	8	AB041749	AB041749 Trichoder	c 605	17	1.5	10686	1	AE005149	AE005149 Halobacte
c 533	17	1.5	2106	9	AB049863	AB049863 Macaca fa	c 606	17	1.5	10729	1	AB012767	AB012767 Pseudomon
c 534	17	1.5	2193	6	AX040670	AX040670 Sequence	c 607	17	1.5	10948	2	AC014767	AC014767 Drosophila
c 535	17	1.5	2282	10	MUSKRF3	D12645 Mouse mRNA	c 608	17	1.5	11113	1	AE007221	AE007221 Sinorhizo
c 536	17	1.5	2298	10	AF454703	AF454703 Mus muscu	c 609	17	1.5	11224	1	AE002229	AE002229 Chlamydo
c 537	17	1.5	2342	10	BC010552	BC010552 Mus muscu	c 610	17	1.5	11230	1	AE001592	AE001592 Chlamydo
c 538	17	1.5	2463	1	BACCGTC	M1880 Bacillus sp	c 611	17	1.5	11250	1	AE009115	AE009115 Agrobacte
c 539	17	1.5	2550	2	AC014237	AC014237 Drosophila	c 612	17	1.5	11300	1	AE002808	AE002808 X98451 Ralstonia s
c 540	17	1.5	2554	10	BC018551	BC018551 Mus muscu	c 613	17	1.5	11374	1	AE008080	AE008080 Agrobacte
c 541	17	1.5	2580	9	SCVERPRL	AF010472 Homo sapi	c 614	17	1.5	11374	1	AE009451	AE009451 Brucella
c 542	17	1.5	2677	8	EO2670	Z26645 S.cerevisia	c 615	17	1.5	12302	1	AE000137	AE000137 Streptoc
c 543	17	1.5	2686	6	EO2161	EO2161 Cycloalide	c 616	17	1.5	12540	1	AE008523	AE008523 Streptoc
c 544	17	1.5	2687	6	AF343071	AF343071 Neurospor	c 617	17	1.5	12855	6	SPAC15E1	AL109770 S.pombe c
c 545	17	1.5	2700	10	AF017182	AF017182 Mus muscu	c 618	17	1.5	13288	8	AX353827	AX353827 Sequence
c 546	17	1.5	2761	10	MMU66141	U66141 Mus muscu	c 619	17	1.5	13323	1	AE004869	AE004869 Pseudomon
c 547	17	1.5	2778	2	AC015234	AC015234 Drosophila	c 620	17	1.5	13398	3	AF346403	AF346403 Eschinococ
c 548	17	1.5	2795	9	S75038	S75038 pancreatic	c 621	17	1.5	14161	1	AE007077	AE007077 Mycobacte
c 549	17	1.5	2863	6	EO2517	EO2517 cDNA encodi	c 622	17	1.5	14734	1	AC104736	AC104736 Pseudomon
c 550	17	1.5	2869	6	AY039981	AY039981 Arabidops	c 623	17	1.5	15501	2	AC108961	AC108961 Rattus no
c 551	17	1.5	2869	6	AR149362	AR149362 Sequence	c 624	17	1.5	15902	1	AC014667	AC014667 Drosophila
c 552	17	1.5	2976	9	AK022892	AK022892 Homo sapi	c 625	17	1.5	16736	2	AF035937	AF035937 Pseudomon
c 553	17	1.5	2994	9	S75037	S75037 pancreatic	c 626	17	1.5	16966	1	CEY55D9A	AL032649 Caenorhab
c 554	17	1.5	2994	9	S75037	S75037 pancreatic	c 627	17	1.5	17509	2	AC109940	AC109940 Rattus no
c 555	17	1.5	3048	1	BACCGTA	M17366 Bacillus sp	c 628	17	1.5	17687	1	AE001038	AE001038 Archaeogl
c 556	17	1.5	3107	2	AC019520	AC019520 Drosophila	c 629	17	1.5	18134	9	AC009775	AC009775 Homo sapi
c 557	17	1.5	3233	3	AE002947	AE002947 Drosophila	c 630	17	1.5	19436	1	U67522	U67522 Methanococ
c 558	17	1.5	3316	6	EO2518	EO2518 cDNA encodi	c 631	17	1.5	22095	2	AC020042	AC020042 Drosophila
c 559	17	1.5	3423	8	SCVRP1GEN	X87806 S.cerevisia	c 632	17	1.5	22504	2	AC109931	AC109931 Rattus no
c 560	17	1.5	3429	6	EO3204	EO3204 DNA encodin	c 633	17	1.5	23304	2	AC109421	AC109421 Rattus no
c 561	17	1.5	3484	6	IO9286	IO9286 Sequence 2	c 634	17	1.5	24629	9	AL627247	AL627247 Human DNA
c 562	17	1.5	3630	6	EO3202	EO3202 DNA encodin	c 635	17	1.5	26104	1	AF052744	AF052744 Yersinia
c 563	17	1.5	3636	6	BC018127	BC018127 Homo sapi	c 636	17	1.5	26201	3	CER24E1	Z49912 Caenorhabdi
c 564	17	1.5	3688	6	EO2516	EO2516 cDNA encodi	c 637	17	1.5	26301	1	AF369954	AF369954 Yersinia
c 565	17	1.5	3724	4	BOVPAMA12	M18683 Bovine PAM	c 638	17	1.5	26378	9	AL645925	AL645925 Human DNA
c 566	17	1.5	3748	8	HUMPAMA12	M37721 Human pepi	c 639	17	1.5	28440	8	SPAC610	Z98603 S.pombe chr
c 567	17	1.5	3822	8	BLVACT1	M56753 Hordeum vul	c 640	17	1.5	29655	2	AC020348	AC020348 Drosophila
c 568	17	1.5	3845	10	EO3203	EO3203 DNA encodin	c 641	17	1.5	30698	2	AC108290	AC108290 Rattus no
c 569	17	1.5	3845	10	AY030276	AY030276 Rattus no	c 642	17	1.5	30829	9	HSCEN95B9	Z69707 Human DNA s
c 570	17	1.5	3913	6	HSMB02146	AL137442 Homo sapi	c 643	17	1.5	31772	2	AC020469	AC020469 Drosophila
c 571	17	1.5	4045	6	EO3201	EO3201 DNA encodin	c 644	17	1.5	33311	9	AC004235	AC004235 Homo sapi
c 572	17	1.5	4046	4	HRSFGML	D29623 Equine mRNA	c 645	17	1.5	33910	9	AP001418	AP001418 Homo sapi
c 573	17	1.5	4140	1	AF302787	AF302787 Bacillus	c 646	17	1.5	34094	14	AB026117	AB026117 Porcine a
c 574	17	1.5	4563	6	AX281721	AX281721 Sequence	c 647	17	1.5	34094	14	AF083132	AF083132 Porcine a
c 575	17	1.5	4803	6	AB024516	AB024516 Colletotr	c 648	17	1.5	34094	14	PAD237815	PAD237815 Streptomy
c 576	17	1.5	4977	6	AX329720	AX329720 Sequence	c 649	17	1.5	34725	1	SC1B2	Z68003 Caenorhabdi
c 577	17	1.5	4977	9	HSU12767	U12767 Human micog	c 650	17	1.5	35409	3	CER02H4	AC004589 Homo sapi
c 578	17	1.5	5066	6	BD003759	BD003759 Polynucle	c 651	17	1.5	35488	9	AC004589	AC004589 Homo sapi
c 579	17	1.5	5119	9	AF333769	AF333769 Homo sapi	c 652	17	1.5	35559	7	PABY13918	Y13918 Pseudomonas
c 580	17	1.5	5164	2	AC013187	AC013187 Drosophila	c 653	17	1.5	35580	7	AB008550	AB008550 Pseudomon
c 581	17	1.5	5889	14	COCPMVB	X00206 Cowpea mosa	c 654	17	1.5	36181	1	U000023	U00023 Mycobacteri
c 582	17	1.5	6327	1	NME391261	AJ391261 Neisseria	c 655	17	1.5	36881	2	AC014041	AC014041 Drosophila
c 583	17	1.5	6638	6	AR036183	AR036183 Sequence	c 656	17	1.5	36882	3	U21324	U21324 Caenorhabdi
c 584	17	1.5	6827	14	AF238884	AF238884 Botrytis	c 657	17	1.5	37269	3	AE002742	AE002742 Drosophila

658	17	1.5	37541	8	YSL8300	U19028 Saccharomyc	731	17	1.5	72709	2	AC026522	AC026522 Homo sapi
659	17	1.5	37644	6	A95302	A95302 Sequence 33	732	17	1.5	72722	2	AC004373	AC004373 Drosophila
660	17	1.5	38086	3	CEFA9B2	Z81543 Caenorhabdi	733	17	1.5	72749	2	AC100270	AC100270 Mus muscu
661	17	1.5	38956	2	AC097983	AC097983 Rattus no	734	17	1.5	72749	2	AC102270	AC102270 Mus muscu
662	17	1.5	39856	9	AF241735	AF241735 Homo sapi	735	17	1.5	72826	2	AL391539	AL391539 Human DNA
663	17	1.5	40351	9	HSU93163	U93163 Homo sapien	736	17	1.5	73061	2	AL102052	AL102052 Mus muscu
664	17	1.5	40619	8	AP003246	AP003246 Oryza sat	737	17	1.5	73357	2	AL590367	AL590367 Human DNA
665	17	1.5	40627	8	HS93B000	AL596189 Homo sapi	738	17	1.5	73858	2	AP001941	AP001941 Homo sapi
666	17	1.5	41494	2	AC017316	AC017316 Drosophila	739	17	1.5	74085	2	AC101421	AC101421 Mus muscu
667	17	1.5	41778	3	AF036700	AF036700 Caenorhab	740	17	1.5	74183	2	AC105624	AC105624 Rattus no
668	17	1.5	42861	1	MTCV339	Z77163 Mycobacteri	741	17	1.5	74658	8	OSR243961	OSR243961 Oryza sat
669	17	1.5	43429	3	CBR644L02	AC084620 Caenorhab	742	17	1.5	75359	8	AL590139	AL590139 Human DNA
670	17	1.5	44495	3	AL589846	AL589846 Human DNA	743	17	1.5	76529	8	HO423H10	HO423H10 Oryza sat
671	17	1.5	45289	3	CBRC39N24	AC084576 Caenorhab	744	17	1.5	78480	2	AC096701	AC096701 Rattus no
672	17	1.5	45773	3	AC109258	AC109258 Mus muscu	745	17	1.5	78765	2	AC021584	AC021584 Homo sapi
673	17	1.5	46239	1	AC027138	AC027138 Staphyloc	746	17	1.5	79676	8	AB013396	AB013396 Arabidops
674	17	1.5	46579	9	AF375663	AF375663 Homo sapi	747	17	1.5	80095	3	AC004283	AC004283 Drosophila
675	17	1.5	46857	2	AC100136	AC100136 Mus muscu	748	17	1.5	80343	3	HS474112	HS474112 Human DNA
676	17	1.5	48152	2	AC096613	AC096613 Rattus no	749	17	1.5	80970	2	AC074227	AC074227 Arabidops
677	17	1.5	48175	2	AC087360	AC087360 Homo sapi	750	17	1.5	81054	2	AL356052	AL356052 Human DNA
678	17	1.5	48974	6	AR161611	AR161611 Sequence	751	17	1.5	81284	8	AB016880	AB016880 Arabidops
679	17	1.5	48974	10	AF007560	AF007560 Mus muscu	752	17	1.5	81624	3	AC099763	AC099763 Caenorhab
680	17	1.5	49162	9	AP001867	AP001867 Homo sapi	753	17	1.5	81673	2	AC107504	AC107504 Rattus no
681	17	1.5	50502	9	HS4494016	AL117328 Human DNA	754	17	1.5	83000	9	AP002960	AP002960 Homo sapi
682	17	1.5	52158	2	AC013185	AC013185 Drosophila	755	17	1.5	83167	9	AL590423	AL590423 Human DNA
683	17	1.5	52970	2	AC108069	AC108069 Homo sapi	756	17	1.5	83711	2	AC099171	AC099171 Rattus no
684	17	1.5	53525	9	AL592437	AL592437 Human DNA	757	17	1.5	86157	2	AC103014	AC103014 Rattus no
685	17	1.5	53585	2	AL357121	AL357121 Homo sapi	758	17	1.5	86535	9	HS491M17	HS491M17 Human DNA
686	17	1.5	53985	9	AL357121	AL357121 Homo sapi	759	17	1.5	86618	9	HS593FE22	HS593FE22 Human DNA
687	17	1.5	54375	2	AC108778	AC108778 Mus muscu	760	17	1.5	87224	9	AC019090	AC019090 Homo sapi
688	17	1.5	54594	2	AC101938	AC101938 Mus muscu	761	17	1.5	88406	10	AF144093	AF144093 Mus muscu
689	17	1.5	55061	2	AC091597	AC091597 Mus muscu	762	17	1.5	88672	2	AC014992	AC014992 Drosophila
690	17	1.5	55361	8	B0311F12	AL512548 Oryza sat	763	17	1.5	88708	3	AC109643	AC109643 Caenorhab
691	17	1.5	55736	2	AC108833	AC108833 Mus muscu	764	17	1.5	89536	2	OSJN00026	OSJN00026 Oryza sat
692	17	1.5	57440	9	AP001828	AP001828 Homo sapi	765	17	1.5	89666	8	AC022403	AC022403 Homo sapi
693	17	1.5	57655	2	AC102327	AC102327 Mus muscu	766	17	1.5	90093	2	AP003241	AP003241 Oryza sat
694	17	1.5	58282	2	AC106326	AC106326 Rattus no	767	17	1.5	90275	2	AP003978	AP003978 Oryza sat
695	17	1.5	58390	2	AC012426	AC012426 Homo sapi	768	17	1.5	91059	9	AP003463	AP003463 Homo sapi
696	17	1.5	59015	2	AC067811	AC067811 Homo sapi	769	17	1.5	92141	9	AL353651	AL353651 Human DNA
697	17	1.5	59652	2	AC084205	AC084205 Homo sapi	770	17	1.5	92698	5	AL136315	AL136315 Human DNA
698	17	1.5	59947	9	AL606833	AL606833 Human DNA	771	17	1.5	92863	5	GGBILOCUS	GGBILOCUS Gallus ga
699	17	1.5	60193	9	AP001425	AP001425 Homo sapi	772	17	1.5	93443	9	HSU63630	HSU63630 Homo sapien
700	17	1.5	62000	9	AP001984	AP001984 Homo sapi	773	17	1.5	93668	9	CNS07EFZ	CNS07EFZ Human chr
701	17	1.5	62685	2	AC100145	AC100145 Mus muscu	774	17	1.5	94074	9	HSBC17A96	HSBC17A96 Homo sapi
702	17	1.5	63441	2	AC100976	AC100976 Mus muscu	775	17	1.5	94459	10	MMTSXDNA	MMTSXDNA M. musculus
703	17	1.5	63618	2	AC110573	AC110573 Mus muscu	776	17	1.5	94829	9	AL357272	AL357272 Homo sapi
704	17	1.5	64196	2	AC067994	AC067994 Homo sapi	777	17	1.5	95214	9	AC012457	AC012457 Homo sapi
705	17	1.5	64387	2	AC091407	AC091407 Rattus no	778	17	1.5	95796	2	AC094350	AC094350 Rattus no
706	17	1.5	64577	3	AE002828	AE002828 Drosophila	779	17	1.5	96146	9	AC095052	AC095052 Homo sapi
707	17	1.5	64786	2	AC017805	AC017805 Drosophila	780	17	1.5	96400	2	AL356516	AL356516 Human DNA
708	17	1.5	64810	9	AL391535	AL391535 Human DNA	781	17	1.5	96565	8	ATT20L15	ATT20L15 Arabidops
709	17	1.5	64975	2	AL391535	AL391535 Human DNA	782	17	1.5	96649	8	AC104283	AC104283 Oryza sat
710	17	1.5	65114	9	AL590225	AL590225 Human DNA	783	17	1.5	97263	8	AC013258	AC013258 Arabidops
711	17	1.5	65130	2	AC026845	AC026845 Homo sapi	784	17	1.5	97817	8	AC090366	AC090366 Homo sapi
712	17	1.5	65955	8	AP003681	AP003681 Oryza sat	785	17	1.5	98259	8	AC002392	AC002392 Arabidops
713	17	1.5	66007	2	AC104039	AC104039 Homo sapi	786	17	1.5	98303	2	AP004642	AP004642 Oryza sat
714	17	1.5	66059	2	AC097999	AC097999 Rattus no	787	17	1.5	98581	8	AL139231	AL139231 Human DNA
715	17	1.5	66315	2	AC095724	AC095724 Rattus no	788	17	1.5	98868	2	AC097797	AC097797 Rattus no
716	17	1.5	66577	2	AC101954	AC101954 Mus muscu	789	17	1.5	98975	2	AL139427	AL139427 Homo sapi
717	17	1.5	66671	2	AC108415	AC108415 Mus muscu	790	17	1.5	98985	2	ATAC012395	ATAC012395 Arabidops
718	17	1.5	67203	2	AC090264	AC090264 Homo sapi	791	17	1.5	99128	8	AC110492	AC110492 Homo sapi
719	17	1.5	67214	2	AC101561	AC101561 Mus muscu	792	17	1.5	99281	2	AC108548	AC108548 Rattus no
720	17	1.5	67214	2	AC101561	AC101561 Mus muscu	793	17	1.5	99539	5	AL591510	AL591510 zebrafish
721	17	1.5	68222	9	AL391904	AL391904 Human DNA	794	17	1.5	99579	9	AP000008	AP000008 Homo sapi
722	17	1.5	68402	9	AL391904	AL391904 Human DNA	795	17	1.5	100000	9	AP000019	AP000019 Homo sapi
723	17	1.5	68631	9	AL139348	AL139348 Human DNA	796	17	1.5	100000	9	AP000070	AP000070 Homo sapi
724	17	1.5	68875	2	AC102330	AC102330 Mus muscu	797	17	1.5	100000	9	AP000019	AP000019 Homo sapi
725	17	1.5	68875	2	AC102330	AC102330 Mus muscu	798	17	1.5	100000	9	AP000019	AP000019 Homo sapi
726	17	1.5	69214	2	AC091182	AC091182 Homo sapi	799	17	1.5	100000	2	AC105363	AC105363 Oryza sat
727	17	1.5	69284	2	AC023185	AC023185 Homo sapi	800	17	1.5	100306	2	AC091810	AC091810 Homo sapi
728	17	1.5	70434	2	AC087266	AC087266 Homo sapi	801	17	1.5	100473	9	HS184J9	HS184J9 Human DNA
729	17	1.5	70981	2	AC101258	AC101258 Mus muscu	802	17	1.5	100951	9	AC103024	AC103024 Rattus no
730	17	1.5	72166	2	AC110517	AC110517 Mus muscu	803	17	1.5	101211	2	AC103024	AC103024 Rattus no

C 804	17	1.5	101248	2	AP003884	AP003884 Oryza sat	C 877	17	1.5	117309	9	AC007274	AC007274 Homo sapi
C 805	17	1.5	101559	2	AC095646	AC095646 Rattus no	C 878	17	1.5	117553	2	AP000595	AP000595 Homo sapi
C 806	17	1.5	101559	2	AC095646	AC095646 Rattus no	C 879	17	1.5	117689	2	AC098671	AC098671 Rattus no
C 807	17	1.5	101579	2	AC008227	AC008227 Drosophi	C 880	17	1.5	117839	8	AC097447	AC097447 Oryza sat
C 808	17	1.5	101648	1	AF339846	AF339846 Microscil	C 881	17	1.5	118103	2	AC009817	AC009817 Homo sapi
C 809	17	1.5	101980	2	AP004007	AP004007 Oryza sat	C 882	17	1.5	118600	2	AC106732	AC106732 Homo sapi
C 810	17	1.5	102266	2	AP003877	AP003877 Oryza sat	C 883	17	1.5	118763	2	AP004228	AP004228 Homo sapi
C 811	17	1.5	103234	2	AP003997	AP003997 Oryza sat	C 884	17	1.5	119201	2	AC109343	AC109343 Homo sapi
C 812	17	1.5	103277	2	AC087824	AC087824 Homo sapi	C 885	17	1.5	119807	2	AP003753	AP003753 Oryza sat
C 813	17	1.5	103417	2	AC094955	AC094955 Rattus no	C 886	17	1.5	119974	2	AP003705	AP003705 Oryza sat
C 814	17	1.5	103573	8	AC005499	AC005499 Arabidops	C 887	17	1.5	120321	2	AP000858	AP000858 Homo sapi
C 815	17	1.5	103933	9	HS684024	HS684024 Arabidops	C 888	17	1.5	120329	2	AP004083	AP004083 Homo sapi
C 816	17	1.5	104389	9	AC090024	AC090024 Homo sapi	C 889	17	1.5	121392	2	AP003824	AP003824 Oryza sat
C 817	17	1.5	104926	9	AC010168	AC010168 Homo sapi	C 890	17	1.5	121392	2	AP003824	AP003824 Oryza sat
C 818	17	1.5	105087	9	AC093884	AC093884 Homo sapi	C 891	17	1.5	121655	2	AC079923	AC079923 Homo sapi
C 819	17	1.5	105360	9	AC093218	AC093218 Homo sapi	C 892	17	1.5	122074	2	AC105927	AC105927 Oryza sat
C 820	17	1.5	105392	9	AC073215	AC073215 Homo sapi	C 893	17	1.5	122074	2	AP004093	AP004093 Oryza sat
C 821	17	1.5	105426	2	AC095315	AC095315 Rattus no	C 894	17	1.5	122383	9	HS7944N18	HS7944N18 Homo sapi
C 822	17	1.5	105538	2	AC104710	AC104710 Oryza sat	C 895	17	1.5	122583	9	AC092435	AC092435 Homo sapi
C 823	17	1.5	105743	8	AC069552	AC069552 Genomic S	C 896	17	1.5	122672	9	AP000704	AP000704 Homo sapi
C 824	17	1.5	105805	2	AC022280	AC022280 Drosophi	C 897	17	1.5	122736	2	AP004080	AP004080 Oryza sat
C 825	17	1.5	105841	8	AC092697	AC092697 Oryza sat	C 898	17	1.5	123196	2	AP004384	AP004384 Oryza sat
C 826	17	1.5	106605	2	AC104055	AC104055 Rattus no	C 899	17	1.5	123427	8	AC091732	AC091732 Oryza sat
C 827	17	1.5	106827	2	AC020829	AC020829 Mus muscu	C 900	17	1.5	123953	2	AP003747	AP003747 Oryza sat
C 828	17	1.5	106827	2	AC004929	AC004929 Homo sapi	C 901	17	1.5	124000	2	AC004060	AC004060 Homo sapi
C 829	17	1.5	107238	9	AC106786	AC106786 Homo sapi	C 902	17	1.5	124129	2	OSJN00097	OSJN00097 Oryza sat
C 830	17	1.5	107397	2	AP000612	AP000612 Homo sapi	C 903	17	1.5	124253	2	AP003990	AP003990 Oryza sat
C 831	17	1.5	107708	2	AL139803	AL139803 Homo sapi	C 904	17	1.5	124309	8	AC091708	AC091708 Gallus ga
C 832	17	1.5	107914	2	AP003949	AP003949 Oryza sat	C 905	17	1.5	124321	8	OSR245900	OSR245900 Oryza sat
C 833	17	1.5	107946	2	AC093945	AC093945 Rattus no	C 906	17	1.5	124351	9	AC005185	AC005185 Homo sapi
C 834	17	1.5	108145	2	AP004649	AP004649 Oryza sat	C 907	17	1.5	124431	2	AC105510	AC105510 Rattus no
C 835	17	1.5	108633	2	AL133518	AL133518 Homo sapi	C 908	17	1.5	124419	2	HSJ558114	HSJ558114 Homo sapi
C 836	17	1.5	108705	2	AP004182	AP004182 Oryza sat	C 909	17	1.5	125823	9	HSJ561124	HSJ561124 Homo sapi
C 837	17	1.5	108904	2	AC103550	AC103550 Oryza sat	C 910	17	1.5	126441	2	AP004151	AP004151 Oryza sat
C 838	17	1.5	108989	2	AP003740	AP003740 Oryza sat	C 911	17	1.5	126482	2	AL137220	AL137220 Homo sapi
C 839	17	1.5	109198	8	AC078891	AC078891 Oryza sat	C 912	17	1.5	126652	2	AL161441	AL161441 Homo sapi
C 840	17	1.5	109398	9	AC044817	AC044817 Homo sapi	C 913	17	1.5	126786	2	OSJN00015	OSJN00015 Homo sapi
C 841	17	1.5	109450	8	AC019498	AC019498 Drosophi	C 914	17	1.5	127203	2	AC106317	AC106317 Rattus no
C 842	17	1.5	109475	8	AP002881	AP002881 Oryza sat	C 915	17	1.5	127295	2	AP002346	AP002346 Homo sapi
C 843	17	1.5	110000	2	AC007194_2	AC007194_2 Continuation (3 of	C 916	17	1.5	127418	9	HS288H20	HS288H20 Homo sapi
C 844	17	1.5	110000	2	AC007194_3	AC007194_3 Continuation (4 of	C 917	17	1.5	127934	2	AC073709	AC073709 Mus muscu
C 845	17	1.5	110000	2	CEY105E8_1	CEY105E8_1 Continuation (2 of	C 918	17	1.5	127934	2	AC073709	AC073709 Mus muscu
C 846	17	1.5	110000	2	CEY105E8_2	CEY105E8_2 Continuation (3 of	C 919	17	1.5	127982	2	AC104473	AC104473 Oryza sat
C 847	17	1.5	110000	2	PFMAL4P1_2	PFMAL4P1_2 Continuation (3 of	C 920	17	1.5	128169	2	HS849L7	HS849L7 Homo sapi
C 848	17	1.5	110042	2	AL590788	AL590788 Human DNA	C 921	17	1.5	128243	2	AC104707	AC104707 Oryza sat
C 849	17	1.5	110042	2	AL590788	AL590788 Human DNA	C 922	17	1.5	128378	2	AC074260	AC074260 Homo sapi
C 850	17	1.5	110074	8	AP003220	AP003220 Oryza sat	C 923	17	1.5	128463	2	AP003619	AP003619 Oryza sat
C 851	17	1.5	110514	8	F7G19	AC000106 Sequence	C 924	17	1.5	128632	2	AC079622	AC079622 Homo sapi
C 852	17	1.5	110637	14	AB024414	AB024414 Gallid he	C 925	17	1.5	128632	8	AP003299	AP003299 Oryza sat
C 853	17	1.5	110775	2	AP003881	AP003881 Oryza sat	C 926	17	1.5	128765	9	AL355852	AL355852 Homo sapi
C 854	17	1.5	110800	8	ATF7K15	AL353871 Arabidops	C 927	17	1.5	128824	1	ECU73857	ECU73857 Homo sapi
C 855	17	1.5	110814	9	AL133342	AL133342 Human DNA	C 928	17	1.5	128903	2	AC093952	AC093952 Oryza sat
C 856	17	1.5	110820	9	AL596285	AL596285 Human DNA	C 929	17	1.5	129090	2	AP003931	AP003931 Oryza sat
C 857	17	1.5	111175	8	AC006585	AC006585 Arabidops	C 930	17	1.5	129118	2	OSJN01000	OSJN01000 Homo sapi
C 858	17	1.5	111271	2	AC095639	AC095639 Rattus no	C 931	17	1.5	129118	2	OSJN01000	OSJN01000 Homo sapi
C 859	17	1.5	111466	2	AC104274	AC104274 Oryza sat	C 932	17	1.5	129242	2	AP004059	AP004059 Oryza sat
C 860	17	1.5	112239	9	AC008648	AC008648 Homo sapi	C 933	17	1.5	129272	2	AL451125	AL451125 Homo sapi
C 861	17	1.5	112664	8	AP003337	AP003337 Oryza sat	C 934	17	1.5	130377	2	OSJN00142	OSJN00142 Oryza sat
C 862	17	1.5	112752	2	AC094583	AC094583 Rattus no	C 935	17	1.5	130451	2	AP004000	AP004000 Oryza sat
C 863	17	1.5	112855	2	AP003608	AP003608 Oryza sat	C 936	17	1.5	130618	2	OSJN00103	OSJN00103 Homo sapi
C 864	17	1.5	112904	9	AC027342	AC027342 Homo sapi	C 937	17	1.5	130630	2	AC010637	AC010637 Homo sapi
C 865	17	1.5	113165	2	AL357112	AL357112 Homo sapi	C 938	17	1.5	130853	5	AL591370	AL591370 Homo sapi
C 866	17	1.5	113720	2	AP003799	AP003799 Oryza sat	C 939	17	1.5	131033	5	AL591370	AL591370 Homo sapi
C 867	17	1.5	114005	9	AC094317	AC094317 Rattus no	C 940	17	1.5	131211	2	AP003810	AP003810 Oryza sat
C 868	17	1.5	114085	2	AC093734	AC093734 Homo sapi	C 941	17	1.5	131287	8	OSB6015	OSB6015 Homo sapi
C 869	17	1.5	115173	2	AC012612	AC012612 Homo sapi	C 942	17	1.5	131480	2	AP004117	AP004117 Oryza sat
C 870	17	1.5	115319	2	AC105724	AC105724 Rattus no	C 943	17	1.5	131820	2	AP003868	AP003868 Homo sapi
C 871	17	1.5	115399	2	AP004322	AP004322 Oryza sat	C 944	17	1.5	131955	2	AC025209	AC025209 Homo sapi
C 872	17	1.5	115454	8	AC022796	AC022796 Homo sapi	C 945	17	1.5	132070	2	AP003663	AP003663 Homo sapi
C 873	17	1.5	115721	8	F25A4	AC008263 Arabidops	C 946	17	1.5	132191	2	AP004741	AP004741 Oryza sat
C 874	17	1.5	115762	2	AP000663	AP000663 Homo sapi	C 947	17	1.5	132244	2	AC093178	AC093178 Oryza sat
C 875	17	1.5	116250	2	AC105563	AC105563 Rattus no	C 948	17	1.5	132268	8	AC099733	AC099733 Oryza sat
C 876	17	1.5	116338	2	AP003940	AP003940 Oryza sat	C 949	17	1.5	132647	8	AC073393	AC073393 Oryza sat





|||||  
Db 541 GGTGGAATGATTAAGAAAGCCGACTCTTTGTACATACGTACCAAACTGGCGGAA 600  
Qy 601 tacatgtgtatgataaacatcatcagcagccagcctcctaatggcccttgagctgc 660  
Db 601 TACATGTGTGATGATTAACCTGATCGATCAGCCAGCTCTTAATGGCGCTGTGAGCTGC 660  
Qy 661 atcgggttcaacagaccgcaaaaaatctgtatgatagcaagctcgtlataacagcaaa 720  
Db 661 ATCGGTTTCAACGAGACCCGAAAAATCTGTATGATGATGAGCAAGCTCCGTATACGCAAA 720  
Qy 721 tattcaacaacacggagcgttaagaagaactgtttagcgccagccagatggcatctcg 780  
Db 721 TATTTCACAAACACGAGGAGTTAGAAAGACTGTTACGCGAGCCGACATGGCATCTCG 780  
Qy 781 ggtgtgcatctactacatgataaacactgtaagaatgatgatgattggttttttgag 840  
Db 781 GGTGTGCGCATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Qy 841 acgttcaactacacgaactctacacgtatctacacgtlataagaatlaacgaactctac 900  
Db 841 ACGTTCAACTACAGCAACTACCGATTTCTACCGTTATGAAGAATATACGACTCTAC 900  
Qy 901 gcatatctgttataaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 960  
Db 901 GCATACATATGTTAAAGTAAAGTACGATGATGATGATGATGATGATGATGATGATG 960  
Qy 961 tccagcgcgttcaacacgaacgaataaactgtgtgtaaaatcatcgagatctgtaaa 1020  
Db 961 TCCAGCGCGTTCAACCCGACGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Qy 1021 actgaaaaaactgataaagcgcagatggtttcgcctgtgtctgtctgtccagcag 1080  
Db 1021 ACTGAAAAACAACATGATAAAGCGGACATGTTGCGCTGTGCTGTGCTGTGCTGTG 1080  
Qy 1081 aacggcgccctaatgataacagctgagctgagctgagctgagctgagctgagctg 1137  
Db 1081 AACGGCGCCCTAATATGAAACACGTGGCGATGATGATGATGATGATGATGATGATG 1137

RESULT 2  
AF361355 630 bp mRNA linear ROD 21-DEC-2001  
LOCUS AF361355  
DEFINITION Rattus norvegicus voltage-dependent calcium channel gamma subunit-like protein mRNA, complete cds.  
ACCESSION AF361355  
VERSION AF361355.1 GI:17974543  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 630)  
AUTHORS Chu, P.-J., Robertson, H.M. and Best, P.M.  
TITLE Calcium channel gamma subunits provide insights into the evolution of this gene family  
JOURNAL Gene 280 (1-2), 37-48 (2001)  
PUBMED 11738816  
REFERENCE 2 (bases 1 to 630)  
AUTHORS Chu, P.-J., Robertson, H. and Best, P.M.  
TITLE Direct Submission  
JOURNAL Submitted (14-MAR-2001) Department of Molecular and Integrative Physiology, University of Illinois, 407 S. Goodwin Ave., Urbana, IL 61801, USA  
FEATURES  
source location/Qualifiers  
1. 630  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
1. 630  
/note="putative; pr"  
/codon\_start=1

/product="voltage-dependent calcium channel gamma subunit-like protein"  
/protein\_id="AAL50050.1"  
/db\_xref="GI:17974544"  
/translation="MTAIGAOAYKLGRPHRSPFESFRTLLIIVCTALAVLSSVS  
ICDGHMLIVERLRLGLWTFCTISNISEPHCLRDLSQAVPYANVMGLARVAAAVY  
AATFLELIVSOVEDVRSRKNVIGSYLLVAFILSSGLTFPIILKNQITLMGF  
TLMFCETPALEFFLNATSGIHINLSRSPAGTLAYRKQYDSTSLI"  
BASE COUNT 102 a 198 c 164 g 166 t  
ORIGIN

Query Match 1.9%; Score 22; DB 10; Length 630;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 935 tcatggttcacccctgattgt 956  
Db 470 TCATGGCTTCACCTGATGTT 491

RESULT 3  
AX079087 2122 bp DNA linear PAT 22-FEB-2001  
LOCUS AX079087  
DEFINITION Sequence 8 from Patent WO0107591.  
ACCESSION AX079087  
VERSION AX079087.1 GI:13158661  
KEYWORDS Pseudomonas sp.  
SOURCE Pseudomonas sp.  
ORGANISM Bacteria; Proteobacteria.  
REFERENCE 1 (bases 1 to 2122)  
AUTHORS Fraser, C.M., Venter, C., Tuemmler, B., Hohelsel, J., Duesterhoeft, A., Hilbert, H., Timms, K.N., Moore, E., Straetz, M. and Helm, S.  
TITLE Data sequences which code export systems  
JOURNAL Patent: WO 0107591-A 8 01-FEB-2001;  
THE INSTITUTE FOR GENOMIC RESEARCH (US)  
FEATURES  
source location/Qualifiers  
1. 2122  
/organism="Pseudomonas sp."  
/db\_xref="taxon:306"  
BASE COUNT 411 a 696 c 705 g 309 t 1 others  
ORIGIN

Query Match 1.9%; Score 22; DB 6; Length 2122;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 399 gctgctgagcctggcggtgag 420  
Db 2068 GCTGCTGAGCCTGGCGGTGAG 2089

RESULT 4  
AC027117 69006 bp DNA linear HTG 23-DEC-2000  
LOCUS AC027117  
DEFINITION Homo sapiens chromosome 8 clone RP11-806011 map 8, LOW-PASS SEQUENCE SAMPLING.  
ACCESSION AC027117  
VERSION AC027117.3 GI:11990735  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 69006)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 8, clone RP11-806011  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 69006)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

TITLE  
JOURNAL  
COMMENT

Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Glinde, S., Goette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kartas, A.,  
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Teste, S., Theodore, J., Tirrell, A., Traversa, M., Trifillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, W.

## Direct Submission

Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 23, 2000 this sequence version replaced gi:7705183.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L8867

Center clone name: 806\_O\_11

\* NOTE: This record contains 83 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
744 843: gap of 100 bp in length  
844 1558: contig of 715 bp in length  
1559 1658: gap of 100 bp  
1659 2397: contig of 739 bp in length  
2398 2497: gap of 100 bp  
2498 3228: contig of 731 bp in length  
3229 3328: gap of 100 bp  
3329 4049: contig of 721 bp in length  
4050 4149: gap of 100 bp  
4150 4893: contig of 744 bp in length  
4894 4993: gap of 100 bp  
4994 5743: contig of 750 bp in length  
5744 5843: gap of 100 bp  
5844 6578: contig of 735 bp in length  
6579 6678: gap of 100 bp  
6679 7414: contig of 736 bp in length  
7415 7514: gap of 100 bp  
7515 8231: contig of 717 bp in length  
8232 8331: gap of 100 bp  
8332 9051: contig of 720 bp in length  
9052 9151: gap of 100 bp  
9152 9867: contig of 716 bp in length  
9868 9967: gap of 100 bp  
9968 10703: contig of 736 bp in length  
10704 10803: gap of 100 bp  
10804 11495: contig of 692 bp in length

11496 11595: gap of 100 bp  
11596 12342: contig of 747 bp in length  
12343 12442: gap of 100 bp  
12443 13166: contig of 724 bp in length  
13167 13266: gap of 100 bp  
13267 14036: contig of 770 bp in length  
14037 14136: gap of 100 bp  
14137 14859: contig of 723 bp in length  
14860 14959: gap of 100 bp  
14960 15714: contig of 755 bp in length  
15715 15814: gap of 100 bp  
15815 16570: contig of 756 bp in length  
16571 16770: gap of 100 bp  
16771 17446: contig of 776 bp in length  
17447 17546: gap of 100 bp  
17547 18277: contig of 731 bp in length  
18278 18377: gap of 100 bp  
18378 19116: contig of 739 bp in length  
19117 19216: gap of 100 bp  
19217 19968: contig of 752 bp in length  
19969 20068: gap of 100 bp  
20069 20806: contig of 738 bp in length  
20807 20906: gap of 100 bp  
20907 21647: contig of 741 bp in length  
21648 21747: gap of 100 bp  
21748 22491: contig of 744 bp in length  
22492 22591: gap of 100 bp  
22592 23348: contig of 757 bp in length  
23349 23448: gap of 100 bp  
23449 24211: contig of 763 bp in length  
24212 24311: gap of 100 bp  
24312 25014: contig of 703 bp in length  
25015 25114: gap of 100 bp  
25115 25835: contig of 721 bp in length  
25836 25935: gap of 100 bp  
25936 26659: contig of 724 bp in length  
26660 26759: gap of 100 bp  
26760 27504: contig of 745 bp in length  
27505 27604: gap of 100 bp  
27605 28302: contig of 698 bp in length  
28303 28402: gap of 100 bp  
28403 29091: contig of 689 bp in length  
29092 29191: gap of 100 bp  
29192 29924: contig of 733 bp in length  
29925 30024: gap of 100 bp  
30025 30777: contig of 753 bp in length  
30778 30877: gap of 100 bp  
30878 31616: contig of 739 bp in length  
31617 31716: gap of 100 bp  
31717 32468: contig of 752 bp in length  
32469 32568: gap of 100 bp  
32569 33288: contig of 720 bp in length  
33289 33388: gap of 100 bp  
33389 34134: contig of 746 bp in length  
34135 34234: gap of 100 bp  
34235 34970: contig of 736 bp in length  
34971 35070: gap of 100 bp  
35071 35795: contig of 725 bp in length  
35796 35895: gap of 100 bp  
35896 36637: contig of 742 bp in length  
36638 36737: gap of 100 bp  
36738 37471: contig of 734 bp in length  
37472 37571: gap of 100 bp  
37572 38294: contig of 723 bp in length  
38295 38394: gap of 100 bp  
38395 39149: contig of 755 bp in length  
39150 39249: gap of 100 bp  
39250 39999: contig of 750 bp in length  
40000 40099: gap of 100 bp  
40100 40844: contig of 745 bp in length  
40845 40944: gap of 100 bp  
40945 41671: contig of 727 bp in length  
41672 41771: gap of 100 bp

```

* 41772 42502: contig of 731 bp in length
* 42503 42602: gap of 100 bp
* 42603 43326: contig of 724 bp in length
* 43327 43426: gap of 100 bp
* 43427 44174: contig of 748 bp in length
* 44175 44274: gap of 100 bp
* 44275 44973: contig of 699 bp in length
* 44974 45073: gap of 100 bp
* 45074 45796: contig of 723 bp in length
* 45797 45896: gap of 100 bp
* 45897 46612: contig of 716 bp in length
* 46613 46712: gap of 100 bp
* 46713 47452: contig of 740 bp in length
* 47453 47552: gap of 100 bp
* 47553 48300: contig of 748 bp in length
* 48301 48400: gap of 100 bp
* 48401 49119: contig of 719 bp in length
* 49120 49219: gap of 100 bp
* 49220 49964: contig of 745 bp in length
* 49965 50064: gap of 100 bp
* 50065 50796: contig of 732 bp in length
* 50797 50896: gap of 100 bp
* 50897 51633: contig of 737 bp in length
* 51634 51733: gap of 100 bp
* 51734 52412: contig of 679 bp in length
* 52413 52512: gap of 100 bp
* 52513 53235: contig of 723 bp in length
* 53236 53335: gap of 100 bp
* 53336 54071: contig of 736 bp in length
* 54072 54171: gap of 100 bp
* 54172 54833: contig of 662 bp in length
* 54834 54933: gap of 100 bp
* 54934 55666: contig of 733 bp in length
* 55667 55766: gap of 100 bp
* 55767 56525: contig of 759 bp in length
* 56526 56625: gap of 100 bp
* 56626 57373: contig of 748 bp in length

```

Query Match 1.9%: Score 22; DB 2; Length 69006;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 gggttactatccgtccctgggtg 189  
 |||||||||||||||||||||  
 Db 40961 GGGTTACTATCCGTCTGGGTG 40982

RESULT 5  
 AC097717/c 142978 bp DNA linear PRI 06-JAN-2002  
 DEFINITION Homo sapiens chromosome 2 clone RP11-650N19, complete sequence.  
 ACCESSION AC097717 AC037452  
 VERSION AC097717.3 GI:18072229  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 142978)  
 Waterston, R.H.  
 The sequence of Homo sapiens clone  
 Unpublished  
 2 (bases 1 to 142978)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (21-OCT-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 3 (bases 1 to 142978)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (06-JAN-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

MO 63108, USA  
 On Jan 6, 2002 this sequence version replaced gi:16756374.  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 Contact: submissions@watson.wustl.edu  
 Project Information  
 Center project name: H\_NH0650N19  
 Drafting center: WIBR  
 -----  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 142978  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-650N19"  
 BASE COUNT 43068 a 28143 c 27230 g 44537 t  
 ORIGIN  
 Query Match 1.9%: Score 22; DB 9; Length 142978;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1017 aaaaactgaacaacatgaat 1038  
 |||||||||||||||||||  
 Db 72024 AAAAACTGAACAACATGAAT 72003

RESULT 6  
 AC012529/c standard; DNA; HTG; 168415 BP.  
 ID AC012529  
 XX AC012529;  
 XX AC012529.2  
 SV  
 XX  
 DT 02-NOV-1999 (Rel. 61, Created)  
 DT 28-MAY-2000 (Rel. 63, Last updated, Version 3)  
 XX  
 DE Homo sapiens chromosome 2 clone RP11-106123 map 2, WORKING DRAFT SEQUENCE.  
 DE 20 unordered pieces.  
 XX  
 KW HTG; HTGS\_DRAFT; HTGS\_PHASE1.  
 XX  
 OS Homo sapiens (human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 XX  
 [1]  
 RP 1-168415  
 RP Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,  
 RL "Homo sapiens chromosome 2, clone RP11-106123";  
 RL Unpublished.  
 XX  
 [2]  
 RP 1-168415  
 RP Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,  
 BA Baldwin J., Batta N., Beckerly R., Boguslavsky L., Boukhgalter B.,  
 RA Brown A., Castle A., Colangelo M., Collins S., Collamore A., Cooke P.,  
 RA DeArrellano K., Dewar K., Domino M., Donelan L., Doyle M., Ferreira P.,  
 RA Fitzhugh W., Forrest C., Funke R., Gage D., Galagan J., Gardyna S.,  
 RA Grant G., Hagos B., Heaford A., Horton L., Howard J.C., Johnson R.,  
 RA Jones C., Kann L., Karatas A., Klein J., Lebecky J., Lien C., Locke K.,  
 RA MacDonald P., Marguis N., McKernan P., McKernan K., McLaughlin J.,  
 RA Meldrum J., Morrow J., Naylor T., Norman C.H., O'Connor T., O'Donnell P.,  
 RA Peterson K., Pollara V., Riley R., Roy A., Santos R., Severy P.,  
 RA Stange-Thomann N., Stojanovic N., Subramanian A., Talamas J., Tesfaye S.,  
 RA Tirrell A., Vassiliev H., Vo A., Wheeler J., Wu X., Wyman D., Ye W.J.,  
 RA Zimmer A., Zody M.;  
 RT

RL Submitted (29-OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,  
 Cambridge, MA 02141, USA  
 CC  
 CC On May 26, 2000 this sequence version replaced gi:6139171.  
 CC All repeats were identified using RepeatMasker:  
 CC Smith, A.F.A. & Green, P. (1996-1997)  
 CC <http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 CC ----- Genome Center  
 CC Center: Whitehead Institute/ MIT Center for Genome Research  
 CC Center code: W1BR  
 CC Web site: <http://www-seq.wi.mit.edu>  
 CC Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 CC ----- Project Information  
 CC Center project name: L1324  
 CC Center clone name: 106.I.23  
 CC ----- Summary Statistics  
 CC Sequencing vector: M13; M77815; 100% of reads  
 CC Chemistry: Dye-terminator-amersham; 3% of reads  
 CC Assembly program: Phrap; version 0.960731  
 CC Consensus quality: 121427 bases at least Q40  
 CC Consensus quality: 145469 bases at least Q30  
 CC Consensus quality: 157834 bases at least Q20  
 CC Insert size: 171000; agarose-fp  
 CC Insert size: 166515; sum-of-contigs  
 CC Quality coverage: 3.3 in Q20 bases; agarose-fp  
 CC Quality coverage: 3.4 in Q20 bases.  
 CC \* NOTE: This is a 'working draft' sequence. It currently  
 CC \* consists of 20 contigs. The true order of the pieces  
 CC \* is not known and their order in this sequence record is  
 CC \* arbitrary. Gaps between the contigs are represented as  
 CC \* runs of N, but the exact sizes of the gaps are unknown.  
 CC \* This record will be updated with the finished sequence  
 CC \* as soon as it is available and the accession number will  
 CC \* be preserved.  
 CC 1 1130: contig of 1130 bp in length  
 CC 1131 1230: gap of 100 bp  
 CC 1231 3206: contig of 1976 bp in length  
 CC 3207 3306: gap of 100 bp  
 CC 3307 4710: contig of 1404 bp in length  
 CC 4711 4810: gap of 100 bp  
 CC 4811 8673: contig of 3863 bp in length  
 CC 8674 8773: gap of 100 bp  
 CC 8774 12291: contig of 3518 bp in length  
 CC 12292 12391: gap of 100 bp  
 CC 12392 15356: contig of 2965 bp in length  
 CC 15357 15456: gap of 100 bp  
 CC 15457 20652: contig of 5196 bp in length  
 CC 20653 20752: gap of 100 bp  
 CC 20753 25969: contig of 5217 bp in length  
 CC 25970 26069: gap of 100 bp  
 CC 26070 29851: contig of 3782 bp in length  
 CC 29852 29951: gap of 100 bp  
 CC 29952 33936: contig of 3985 bp in length  
 CC 33937 34036: gap of 100 bp  
 CC 34037 39264: contig of 5228 bp in length  
 CC 39265 39364: gap of 100 bp  
 CC 39365 45952: contig of 6588 bp in length  
 CC 45953 46052: gap of 100 bp  
 CC 46053 52806: contig of 6754 bp in length  
 CC 52807 52906: gap of 100 bp  
 CC 52907 62575: contig of 9669 bp in length  
 CC 62576 62675: gap of 100 bp  
 CC 62676 70800: contig of 8125 bp in length  
 CC 70801 70900: gap of 100 bp  
 CC 70901 82081: contig of 11181 bp in length  
 CC 82082 82181: gap of 100 bp  
 CC 82182 96270: contig of 14089 bp in length  
 CC 96271 96370: gap of 100 bp  
 CC 96371 111552: contig of 15182 bp in length  
 CC 111553 111652: gap of 100 bp  
 CC 111653 135717: contig of 24065 bp in length

CC \* 135718 135817: gap of 100 bp  
 CC \* 135818 168415: contig of 32598 bp in length.  
 CC  
 CC Key  
 CC source  
 CC location/Qualifiers  
 CC 1. 168415  
 CC /chromosome="2"  
 CC /db\_xref="taxon:9606"  
 CC /organism="Homo sapiens"  
 CC /map="2"  
 CC /clone="RP11-106I23"  
 CC /clone\_11p="RP11-11 Human Male BAC"  
 CC 1. 1130  
 CC /note="assembly\_fragment"  
 CC 1231. 3206  
 CC /note="assembly\_fragment"  
 CC 3307. 4710  
 CC /note="assembly\_fragment clone\_end:T7 vector\_side:right"  
 CC 4811. 8673  
 CC /note="assembly\_fragment"  
 CC 8774. 12291  
 CC /note="assembly\_fragment"  
 CC 12392. 15356  
 CC /note="assembly\_fragment"  
 CC 15457. 20652  
 CC /note="assembly\_fragment"  
 CC 20753. 25969  
 CC /note="assembly\_fragment"  
 CC 26070. 29851  
 CC /note="assembly\_fragment"  
 CC 29952. 33936  
 CC /note="assembly\_fragment"  
 CC 34037. 39264  
 CC /note="assembly\_fragment"  
 CC 39365. 45952  
 CC /note="assembly\_fragment"  
 CC 46053. 52806  
 CC /note="assembly\_fragment"  
 CC 52907. 62575  
 CC /note="assembly\_fragment"  
 CC 62676. 70800  
 CC /note="assembly\_fragment clone\_end:SP6 vector\_side:right"  
 CC 70901. 82081  
 CC /note="assembly\_fragment"  
 CC 82182. 96270  
 CC /note="assembly\_fragment"  
 CC 96371. 111552  
 CC /note="assembly\_fragment"  
 CC 111653. 135717  
 CC /note="assembly\_fragment"  
 CC 135818. 168415  
 CC /note="assembly\_fragment"  
 CC  
 CC Sequence 168415 BP; 51193 A; 32641 C; 32171 G; 50421 T; 1989 other;  
 CC  
 CC Query Match 1.9%; Score 22; DB 30; Length 168415;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7;  
 CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1017 aaacatgaataaactgaat 1038  
 CC |  
 CC Db 2104 AAAAAGTGAACAACTGAAT 2083  
 CC  
 CC RESULT 7  
 CC AC108032 173998 bp DNA linear HTG 24-JAN-2002  
 CC LOCUS  
 CC DEFINITION Homo sapiens chromosome 2 clone RP11-106I23, WORKING DRAFT  
 CC SEQUENCE 2 unordered pieces.  
 CC ACCESSION AC108032 AC012529  
 CC VERSION AC108032.1 GI:18308800  
 CC KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS 1 (bases 1 to 173998)  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 173998)  
AUTHORS Waterston, R.H.  
TITLE Unpublished  
JOURNAL Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
COMMENT On Jan 24, 2002 this sequence version replaced gi:8096824.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
Project Information -----  
Center project name: H\_NH0106123  
Drafting center: MIBR

----- Summary Statistics -----  
Sequencing vector: M13; %  
Sequencing vector: plasmid; %  
Chemistry: Dye-primer ET; % of reads  
Chemistry: Dye-terminator Big Dye; % of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171517 bases at least Q40  
Consensus quality: 172379 bases at least Q30  
Consensus quality: 173864 bases at least Q20  
Insert size: 174000; agarose-fp  
Insert size: 175963; sum-of-coverage  
Quality coverage: 7.60 in Q20 bases; agarose-fp  
Quality coverage: 7.52 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 61739: contig of 61739 bp in length  
\* 61740 61839: gap of unknown length  
\* 61840 173998: contig of 112159 bp in length.  
Location/Qualifiers  
1..173998  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/clone="RP11-106123"  
1..61739  
/note="assembly\_name:Contig26  
clone\_end:SP6  
vector\_side:left"  
61840..173998  
/note="assembly\_name:Contig27  
vector\_end:R7  
clone\_end:left"  
misc\_feature  
misc\_feature  
misc\_feature  
BASE COUNT 53651 a 33484 c 34031 g 52732 t 100 others  
ORIGIN  
Query Match 1.9%; Score 22; DB 2; Length 173998;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Gy 1017 aaaactgaaacaactgaaat 1038

Db 103229 AAAAAGTGAACAACTGAAT 103250  
|||||  
RESULT 8  
AC108190  
LOCUS  
DEFINITION  
ACCESSION  
AC108190  
VERSION  
AC108190.1 GI:18376889  
KEYWORDS  
HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
SOURCE  
cat.  
ORGANISM  
Felis catus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
REFERENCE 1 (bases 1 to 132388)  
AUTHORS Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Ho, S.-L., Idol, J.R., Karling, E., Lalic, P., Lee-Lin, S.-O.,  
Legaspi, R., Maduro, Q.L., Maduro, V.B., Masfello, C., Maskeri, B.,  
Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,  
Stantrop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W.,  
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H., and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 132388)  
Green, E.D.  
Direct Submission  
Submitted (26-JAN-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
COMMENT ----- Genome Center -----  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc\_mouse@ngri.nih.gov  
Project Information  
Center project name: cod  
Center clone name: 443102

----- Summary Statistics -----  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 132360 bases at least Q40  
Consensus quality: 132380 bases at least Q30  
Consensus quality: 132388 bases at least Q20  
Insert size: 134000; agarose-fp  
Insert size: 132388; sum-of-coverage  
Quality coverage: 11.04x in Q20 bases; agarose-fp  
Quality coverage: 11.18x in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1 132388: contig of 132388 bp in length.  
Location/Qualifiers  
1..132388  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/clone="RP86-44312"  
/clone\_id="RP86"  
1..132388  
/note="assembly-fragment  
clone\_end:SP6  
vector\_side:left"

FEATURES  
source  
misc\_feature

BASE COUNT 42519 a 25354 c 24854 g 39661 t  
ORIGIN  
Query Match 1.88; Score 21; DB 2; Length 132388;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1016 taaactgaacaaactga 1036  
Db 33965 TAAAACTGAACAACTGA 33985  
RESULT 9  
AC097081 138247 bp DNA linear HTG 20-DEC-2001  
LOCUS AC097081.3  
DEFINITION Rattus norvegicus clone CH230-49G13, \*\*\* SEQUENCING IN PROGRESS  
ACCESSION AC097081.3 GI:17973481  
VERSION AC097081.3  
KEYWORDS HTG: HTGS PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 138247)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C.,  
Aisbrook,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbata,D.,  
Benton,D., Bimge,K., Blankenhug,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brilewa,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Hollway,C.,  
Hollins,B., Homsl,E., Howard,S., Huber,J., Huliy,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W.,  
Louisege,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,  
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,J., Mottabatt,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
Ogun,M., Okwuon,G., Otaguine,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,L., Primus,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,  
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoochhari,N.,  
Stinson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tameris,A., Tameris,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
Worley,K., Wu,C., Wu,X., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (09-OCT-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

## COMMENT

Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:17064380.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GHXN  
Center clone name: CH230-49G13  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329first call to  
findPhrapList  
Consensus quality: 110444 bases at least Q40  
Consensus quality: 117693 bases at least Q30  
Consensus quality: 124112 bases at least Q20  
Estimated insert size: 111960; sum-of-coverage  
Quality coverage: 0x in Q20 bases; agarose-1p estimation  
Quality coverage: 1.5x in Q20 bases; sum-of-coverage  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 59 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
7367: contig of 7367 bp in length  
7368: gap of unknown length  
7468: contig of 6423 bp in length  
13891: gap of unknown length  
13891: contig of 4770 bp in length  
18961: gap of unknown length  
18961: contig of 4273 bp in length  
23334: gap of unknown length  
23334: contig of 2448 bp in length  
25882: gap of unknown length  
25882: contig of 4440 bp in length  
30422: gap of unknown length  
30522: contig of 2174 bp in length  
32695: gap of unknown length  
32795: contig of 3475 bp in length  
36370: gap of unknown length  
36370: contig of 3847 bp in length  
40217: gap of unknown length  
40217: contig of 2977 bp in length  
40318: gap of unknown length  
40318: contig of 2386 bp in length  
42903: gap of unknown length  
42903: contig of 3228 bp in length  
43004: gap of unknown length  
43004: contig of 2977 bp in length  
46332: gap of unknown length  
46332: contig of 2189 bp in length  
49408: gap of unknown length  
49408: contig of 1941 bp in length  
51557: gap of unknown length  
51557: contig of 2049 bp in length  
53950: gap of unknown length  
53950: contig of 2393 bp in length  
53951: gap of unknown length  
53951: contig of 2189 bp in length  
56239: gap of unknown length  
56239: contig of 1941 bp in length  
58280: gap of unknown length  
58280: contig of 2588 bp in length  
60968: gap of unknown length  
60968: contig of 2947 bp in length  
64015: gap of unknown length  
64015: contig of 3094 bp in length  
67209: gap of unknown length  
67209: contig of 2649 bp in length  
69958: gap of unknown length  
69958: contig of 1480 bp in length  
70059: gap of unknown length  
70059: contig of 1789 bp in length  
71538: gap of unknown length  
71538: contig of 1789 bp in length  
71639: gap of unknown length  
71639: contig of 1789 bp in length

```

* 73428 73527: gap of unknown length
* 73528 76068: contig of 2541 bp in length
* 76069 76168: gap of unknown length
* 76169 79477: contig of 3309 bp in length
* 79478 79577: gap of unknown length
* 79578 81785: contig of 2208 bp in length
* 81786 81885: gap of unknown length
* 81886 83753: contig of 1868 bp in length
* 83754 83853: gap of unknown length
* 83854 86171: contig of 2318 bp in length
* 86172 86271: gap of unknown length
* 86272 87499: contig of 1228 bp in length
* 87500 87599: gap of unknown length
* 87600 89038: contig of 1439 bp in length
* 89039 89138: gap of unknown length
* 89139 92546: contig of 3408 bp in length
* 92547 92646: gap of unknown length
* 92647 95159: contig of 2513 bp in length
* 95160 95259: gap of unknown length
* 95260 96596: contig of 1337 bp in length
* 96597 96696: gap of unknown length
* 96697 98116: contig of 1420 bp in length
* 98117 98216: gap of unknown length
* 98217 99622: contig of 1406 bp in length
* 99623 99722: gap of unknown length
* 99723 101221: contig of 1499 bp in length
* 101222 101321: gap of unknown length
* 101322 102454: contig of 1133 bp in length
* 102455 102554: gap of unknown length
* 102555 103839: contig of 1285 bp in length
* 103840 103939: gap of unknown length
* 103940 105230: contig of 1291 bp in length
* 105231 105330: gap of unknown length
* 105331 107000: contig of 1670 bp in length
* 107001 107100: gap of unknown length
* 107101 108797: contig of 1697 bp in length
* 108798 108897: gap of unknown length
* 108898 110317: contig of 1420 bp in length
* 110318 110417: gap of unknown length
* 110418 112498: contig of 2081 bp in length
* 112499 112598: gap of unknown length
* 112599 114388: contig of 1790 bp in length
* 114389 114488: gap of unknown length
* 114489 115691: contig of 1203 bp in length
* 115692 115791: gap of unknown length
* 115792 117646: contig of 1855 bp in length
* 117647 117746: gap of unknown length
* 117747 118865: contig of 1119 bp in length
* 118866 118965: gap of unknown length
* 118966 120061: contig of 1096 bp in length
* 120062 120161: gap of unknown length
* 120162 121287: contig of 1126 bp in length
* 121288 121387: gap of unknown length
* 121388 123004: contig of 1617 bp in length
* 123005 123104: gap of unknown length
* 123105 124973: contig of 1859 bp in length
* 124974 125073: gap of unknown length
* 125074 126303: contig of 1230 bp in length
* 126304 126403: gap of unknown length
* 126404 127649: contig of 1246 bp in length
* 127650 127749: gap of unknown length
* 127750 129482: contig of 1733 bp in length
* 129483 129582: gap of unknown length
* 129583 130645: contig of 1063 bp in length
* 130646 130745: gap of unknown length
* 130746 132461: contig of 1716 bp in length
* 132462 132561: gap of unknown length
* 132562 134179: contig of 1618 bp in length
* 134180 134279: gap of unknown length

```

Query Match 1.8% Score 21: DB 2: Length 138247;  
 Best Local Similarity 100.0% Pred. No. 6.2;  
 Matches 21: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

Oy 311 aaagcctgattgacctgagct 331
|||||
Db 14375 AAAGCCTGATTGACCTGAGCT 14355

RESULT 10
AC108193 139411 bp DNA linear HTG 26-JAN-2002
AC108193/C Felis catus clone RP86-469M8, WORKING DRAFT SEQUENCE, 2 unordered
LOCUS pieces.
DEFINITION
ACCESSION AC108193.1 GI:18376892
VERSION AC108193
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE cat.
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
1 (bases 1 to 139411)
Akher, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Ho, S.-L., Idol, J.R., Karlins, E., Latic, P., Lee-Jin, S.-O.,
Legsapi, R., Maduro, Q.L., Maduro, V.B., Mastiello, C., Master, B.,
Mastrian, S.D., McCloskey, J.C., McDowell, V., Pearson, R., Prasad, A.,
Stantiripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W.,
Tsurgoev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 139411)
Green, E.D.
Direct Submission
Submitted (26-JAN-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nih.gov
Project information
Center project name: coe
Center clone name: 469M08
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 138590 bases at least Q40
Consensus quality: 138720 bases at least Q30
Consensus quality: 138869 bases at least Q20
Insert size: 133000; agarose-fp
Insert size: 139311; sum-of-contigs
Quality coverage: 12.22x in Q20 bases; agarose-fp
Quality coverage: 11.66x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 33148: contig of 33148 bp in length
* 33149 33248: gap of unknown length
* 33249 139411: contig of 106163 bp in length.
FEATURES
source
Location/Qualifiers
1..139411
/organism="Felis catus"
/db_xref="taxon:9685"
/clone_lib="RP86-469M8"
/clone_lib="RP86"
1..33148
misc_feature

```

```

/note="assembly-fragment
clone_end:T7
vector_side:left"
misc_feature 33249..139411
/note="assembly-fragment
clone_end:SP6
vector_side:left"
BASE COUNT 42075 a 26701 c 26611 g 43924 t 100 others
ORIGIN

```

```

Query Match 1.8%; Score 21; DB 2; Length 139411;
Best Local Similarly 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1016 taaaaactgaacaaactga 1036
|||||
Db 60661 TAAAACTGAACAACTGA 60641

```

```

RESULT 11
AC004485 185516 bp. DNA linear PRI 03-FEB-2000
LOCUS AC004485/c
DEFINITION Homo sapiens BAC clone CTB-118E13 from 7p15-p21, complete sequence.
ACCESSION AC004485
VERSION AC004485.1 GI:2992497
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 185516)
AUTHORS Lacy/M.
TITLE The sequence of Homo sapiens BAC clone CTB-118E13
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 185516)
TITLE Waterston,R.
JOURNAL Direct Submission
AUTHORS Submitted (27-MAR-1998) Department of Genetics, Washington
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 3 (bases 1 to 185516)
AUTHORS Waterston,R.
JOURNAL Direct Submission
AUTHORS Submitted (03-FEB-2000) Department of Genetics, Washington
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_Rc118E13

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see

<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send  
mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
Clone CTB-118E13 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-T. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelobAC11  
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
The actual start of this clone is at base position 1 of CTB-118E13; actual end is at 185516 of CTB-118E13. The orientation of this clone is unknown.

This clone contains STS SWS51475 (MID:91222787).

```

FEATURES
    source
        location/qualifiers
            1..185516
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7p15-p21"
            /clone="CTB-118E13"
            /clone_1kb="CTB-978SK-B"
            1376..2438
            /rpt_family="L1"
            repeat_region
                3388..3454
                /rpt_family="MER1_type"
            4144..4443
            /rpt_family="Alu"
            4454..4792
            /rpt_family="L2"
            5182..5483
            /rpt_family="Alu"
            6636..6800
            /rpt_family="MER1_type"
            7781..7854
            /rpt_family="Retroviral"
            7924..8143
            /rpt_family="Retroviral"
            9235..9285
            /rpt_family="MER1"
            10636..10813
            /rpt_family="L1"
            11664..11841
            /rpt_family="MIR"
            13662..13846
            /rpt_family="L1"
            14191..14355
            /rpt_family="MIR"
            14838..14948
            /rpt_family="MER81"
            15640..15820
            /rpt_family="L2"
            17476..17550
            /rpt_family="Mariner"
            17590..17892
            /rpt_family="Alu"
            18483..18564
            /rpt_family="L2"
            18690..18980
            /rpt_family="Alu"
            19570..19699
            /rpt_family="L2"
            20714..20817
            /rpt_family="MIR"
            21043..21384
            /rpt_family="MER2_type"
            22077..22266
            /rpt_family="MER1_type"

```



```

misc_feature      22605..22805
                  /note="match to EST AA502194 (MID:g2237161) ng73e12.sl"
misc_feature      /complement(22605..22805)
misc_feature      /note="match to EST AA226171 (MID:g1847487) nc11g07.r1"
                  /complement(22609..22808)
gene              /note="match to EST AA484674 (MID:g2211468) ne64a03.sl"
                  /complement(22780..29226)
                  /gene="NPY"
CDS               complement(join(22780..22804,24888..24968,29039..29226))
                  /gene="NPY"
                  /note="match to protein P01303 (PID:g128117):
                  H_RG118R13.1"
                  /codon_start=1
                  /product="neuropeptide Y precursor"
                  /protein_id="AAC08287.1"
                  /db_xref="gi:2992498"
                  /translation="MLGNKRRLGSLGTLALSLVLCALAEAPSPKPNPGDAPAD
MARRYSAIRHYINILTRQRYGKRSPETLIDLMRESTENVPRIRLEDPAWM"
repeat_region     23527..23690
                  /rpt_family="Alu"
misc_feature      24859..24970
                  /note="match to EST AA662702 (MID:g2616693) nv06h11.sl"
                  /complement(24859..24970)
misc_feature      /gene="NPY"
                  /note="match to EST AA226171 (MID:g1847487) nc11g07.r1"
                  /note="match to EST AA502194 (MID:g2237161) ng73e12.sl"
repeat_region     26493..26667
                  /rpt_family="MERL_type"
repeat_region     26686..26872
                  /rpt_family="MERL_type"
repeat_region     28886..28964
                  /rpt_family="L2"
misc_feature      29006..30527
                  /note="Cpg_island (%GC=65.4, c/e=0.80, #cpgs=123)"
                  /note="match to EST AA662702 (MID:g2616693) nv06h11.sl"
misc_feature      32483..33035
                  /rpt_family="L1"
repeat_region     33034..33965
                  /rpt_family="L1"
repeat_region     34024..34394
                  /rpt_family="MALR"
repeat_region     34395..34584
                  /rpt_family="L1"
repeat_region     35069..35311
                  /rpt_family="L2"
repeat_region     35345..41477
                  /rpt_family="L1"
repeat_region     41729..42145
                  /rpt_family="MALR"
repeat_region     42234..42446
                  /rpt_family="MERL_type"
repeat_region     42807..42990
                  /rpt_family="L2"
repeat_region     43879..44184
                  /rpt_family="Alu"
repeat_region     44664..44934
                  /rpt_family="Alu"
repeat_region     45003..45251
                  /rpt_family="L1"
repeat_region     45338..45585
                  /rpt_family="L2"
repeat_region     45623..46229
                  /rpt_family="L2"
repeat_region     46230..46277
                  /rpt_family="MER3-group"
repeat_region     46278..46504
                  /rpt_family="L2"
repeat_region     46517..46674
                  /rpt_family="L1"
repeat_region     46715..46852
                  /rpt_family="MERL_type"

```

```

repeat_region     47387..47478
                  /rpt_family="L1"
repeat_region     47824..48086
Query Match      1.8%: Score 21; DB 9; Length 185516;
Best Local Similarity 100.0%; Pred No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 794 accgatgacacacgcgaag 814
DB 49017 ACCTGATGAACACACCTGGAAG 48997
RESULT 12
AL596170 313450 bp DNA linear BCT 04-DBC-2001
LOCUS Listeria innocua Clp11262 complete genome, segment 8/12.
DEFINITION AL596170 AL592022
ACCESSION AL596170.1 GI:16414292
VERSION
KEYWORDS
SOURCE
ORGANISM
Listeria innocua.
Listeria innocua
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Listeria.
1 (sites)
REFERENCE
AUTHORS Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
Bagero,F., Berche,P., Bloeker,H., Brandt,P., Chakraborty,T.,
Charbit,A., Chetoui,F., Couve,E., de Daruvar,A., Deloux,P.,
Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
Dussurget,O., Entian,K.D., Esnault,H., Portillo,F.G., Garrido,P.,
Gautier,L., Goebel,N., Gomez-Lopez,W., Hain,T., Haefl,J.,
Jackson,D., Jones,L.M., Kaerst,U., Kreft,J., Kuhn,M., Kunst,F.,
Kurapat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
Nedjar,H., Nordstedt,G., Novella,S., de Pablo,B., Perez-Diaz,J.C.,
Purcell,R., Remmel,B., Rose,M., Schleuter,T., Simoes,N.,
Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,U. and Cossart,P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)
21537279
2 (bases 1 to 313450)
Glaser,P., Frangeul,L. and Rusniok,C.
Direct Submission
Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
COMMENT
E-mail: pglaser@pasteur.fr
Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES
Source
1..313450
location/Qualifiers
1..313450
/organism="Listeria innocua"
/strain="Clp11262"
/db_xref="taxon:1642"
complement(287..1444)
/gene="lin1790"
complement(287..1444)
/gene="lin1790"
/note="similar to transmembrane transport proteins"
/codon_start=1
/protein_id="CAC97021.1"
/db_xref="gi:16414293"
/translation="MGAKGKFWILTMVAISGLSGVLLPLAIIIEKVGAGNGF
HATGIYGLVAVSPTEAPLHRYGKPIILVGGGLVAIAIFAPITWNLVFWIILKL
IGIGDHMLHFSQWIGAMSDPSKGRNMAIYGFVSQFAIPOLVNLAKVNNAPF
FLSGILVLIAMSLVWFIRNDEYGEKAVIRKISFGSIRKFSGVKAMVAMAPPELYG
TLERGLNAPFPVVGKRDGIDPMIAMIISSEVGTITFOVPIGIIISDKGRDKMLPL
TGAGAVVPLAFVFKPIPEFVIFPFIIGISLISIGSLGSLVWTDIPLLELAGNIV
GMFSLGSIIGSATGMIGIGTGNOIFTFVVGIIILGCLLALGKRAH"
complement(1508..1579)
/note="CRNscan-SE vs 1.3 result - Cove score = 69.21"
/product="tRNA-Glu"
complement(1585..1675)
/note="CRNscan-SE vs 1.3 result - Cove score = 63.76"
tRNA
tRNA

```

tRNA	/product="tRNA-Ser" complement(1712..1784) /note="tRNAscan-SE vs 1.3 result - Cove score = 85.46"
tRNA	/product="tRNA-Asn" complement(1791..1864) /note="tRNAscan-SE vs 1.3 result - Cove score = 92.91"
tRNA	/product="tRNA-Ile" complement(1891..1961) /note="tRNAscan-SE vs 1.3 result - Cove score = 73.1"
tRNA	/product="tRNA-Gly" complement(1984..2056) /note="tRNAscan-SE vs 1.3 result - Cove score = 72.11"
tRNA	/product="tRNA-His" complement(2083..2155) /note="tRNAscan-SE vs 1.3 result - Cove score = 79.39"
tRNA	/product="tRNA-Phe" complement(2160..2235) /note="tRNAscan-SE vs 1.3 result - Cove score = 73.32"
tRNA	/product="tRNA-Asp" complement(2240..2313) /note="tRNAscan-SE vs 1.3 result - Cove score = 78.6"
tRNA	/product="tRNA-Met" complement(2356..2445) /note="tRNAscan-SE vs 1.3 result - Cove score = 66.89"
tRNA	/product="tRNA-Ser" complement(2468..2541) /note="tRNAscan-SE vs 1.3 result - Cove score = 90.88"
tRNA	/product="tRNA-Met" complement(2553..2626) /note="tRNAscan-SE vs 1.3 result - Cove score = 77.07"
tRNA	/product="tRNA-Met" complement(2681..2753) /note="tRNAscan-SE vs 1.3 result - Cove score = 84.64"
tRNA	/product="tRNA-Ala" complement(2773..2846) /note="tRNAscan-SE vs 1.3 result - Cove score = 86.85"
tRNA	/product="tRNA-Pro" complement(2857..2930) /note="tRNAscan-SE vs 1.3 result - Cove score = 74.21"
tRNA	/product="tRNA-Arg" complement(2940..3025) /note="tRNAscan-SE vs 1.3 result - Cove score = 66.88"
tRNA	/product="tRNA-Leu" complement(3040..3114) /note="tRNAscan-SE vs 1.3 result - Cove score = 87.88"
tRNA	/product="tRNA-Gly" complement(3130..3211) /note="tRNAscan-SE vs 1.3 result - Cove score = 69.82"
tRNA	/product="tRNA-Leu" complement(3217..3289) /note="tRNAscan-SE vs 1.3 result - Cove score = 88.37"
tRNA	/product="tRNA-Lys" complement(3330..3405) /note="tRNAscan-SE vs 1.3 result - Cove score = 95.38"
tRNA	/product="tRNA-Thr" complement(3414..3486) /note="tRNAscan-SE vs 1.3 result - Cove score = 86.53"
tRNA	/product="tRNA-Val" complement(3503..3612) /gene="tRNA-5s"
gene	complement(3503..3612) /gene="tRNA-5s"
tRNA	complement(3693..6623) /gene="tRNA-23s"
gene	complement(3693..6623) /gene="tRNA-23s"
tRNA	complement(6669..8423) /gene="tRNA-16s"
gene	complement(6669..8423) /gene="tRNA-16s"
terminator	complement(8850..8868) /gene="tRNA-16s"
gene	complement(8850..8868) /gene="tRNA-16s"
gene	complement(8850..9335) /gene="tRNA-16s"

CDS	/gene="lin1791" complement(876..9322) /gene="lin1791" /note="similar to transcription regulators (fur family), perr in B. subtilis" /codon_start=1 /transl_table=1 /protein_id="CAC97022.1" /db_xref="GI:16414294"
RBS	/translation="MAVSNAITKEAVDVILKGTGVRITPORHATIEFLINSHPTPAD IYRSLGEPNMSATYVNNIRVVRDGLIKELSGASSRPPFSNTHALCNWGS KIVDFHRTGDEVEHFAHVTGYEIDNHRLVEYGTCECEKOSNN"
gene	complement(9330..9335) /gene="lin1791"
CDS	complement(9405..10361) /gene="lin1792" complement(9405..10346) /gene="lin1792" /note="similar to glycinate dehydrogenases" /codon_start=1 /transl_table=1 /protein_id="CAC97023.1" /db_xref="GI:16414295"
gene	/translation="MKYLFILDPPEERKKNQAEKPPDPTFESIDHFNUGEVYII TYGNSITTEKIKKATNLKLTIVSAGVDSIPTKIDOKITVANVGHAIIPWGETAL SEFMLSHVKKAFYKMQEKEMASEEPTILACTIVAGTGAIGAKVAEFAQFME VIGINTGHVPKPFSTYAMTDLEKVAFLDFVSYLPQTEESGTYOLSEFEKMTN AVFINIGRSVLEETLERASKEQIAHFYLDVPEPLPAESLYMQASVNTTTPHVS GHSKDYDRSEFWFENINHLKENTKLRNEDLNGKY"
RBS	complement(10356..10361) /gene="lin1792"
gene	complement(10476..10481) /gene="gsab"
RBS	10476..10481 /gene="gsab"
CDS	10489..11787 /gene="gsab" /codon_start=1 /transl_table=1 /product="glutamate-1-semialdehyde aminotransferase" /protein_id="CAC97024.1"
gene	/db_xref="GI:16414296"
RBS	/translation="MDHSMKSKLHDEALLHIIGGVNSPSRSKNGVGGIPVTERANG AYFYVDGNKYIDILAAFGPIITGHAPHITTAITNAONGVYGPTRHETFPAML KEALPSLEKRYFTSGTEAVMTITRVAVTGDKLIKFGAGYHGHFDVLYAGSGP STGIDPSAGVTRSTAEVITVPNDLTSKALAWGDOVAIVLEVTGNGVAP AEGFLAVNELAHNGSLVYDVTAPFRMNGCAQNYGVLDLTMKRTIIGCGPI GATGGRVDIMKVAPLCPAYOAGTHAGNPASITSGTCEVLOEBGLYDFREKYSML KEGLEKAIKHGIAVTVNQIVGALTYETDEPVTNAEAGATGDLFGRPFGMLEEG INLAPSKYEAWFITSAHSEADILETIOAVDTVGKLVQGR"
RBS	11879..11884 /gene="lin1794"
gene	join(11879..12981,12990..13016) /gene="lin1794"
CDS	11893..12981 /gene="lin1794" /note="similar to hypothetical proteins" /codon_start=1 /transl_table=1 /protein_id="CAC97025.1" /db_xref="GI:16414297"
gene	/translation="MKFGARILKTYGIAITLAEFIAQLCNPSPSLAGISAVPAIOPSI YRSYRTILERAQGNVIGALITAEFGIYIGNDFTLIVASICVALLMOPRLBNTIGLA VVLITVMDSPGNDPLFIALIRFGTIMGILAAFIYNLFLPKYEVSLFOIYVNS EIVRWIKINLRHAADFPILKKDMEMKOLNTRNLVGLYREBRFLKKNALSKGRKI AVYRQMLCSQKGFELIKIQRHYENDYLDLPKPKGLINQHIDYLDKREOLLTYID KVSIDLEYESHLDQPDQMLFLREMEETEDVEMDMKTHLMKRTIASVFAIOET IDYLERKIHSEFLRHTENQIDINVEE"
gene	complement(join(12990..13016,13022..13549,13557..13562))

Query Match 1.88; Score 21; DB 1; Length 313450;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 acccgtaaggatcatcca 167  
|||||  
Db 241631 ATCCGTAAGGTATCATCCA 241651

RESULT 13  
AL591981 347050 bp DNA linear BCT 04-DEC-2001  
LOCUS Listeria monocytogenes strain EGD, complete genome, segment 9/12.  
DEFINITION AL591981 AL591824  
ACCESSION AL591981.1 GI:16411141  
VERSION  
KEYWORDS  
SOURCE Listeria monocytogenes.  
ORGANISM Listeria monocytogenes.  
Bacteria; Firmicutes; Bacillus/Clostridium group;  
Bacillus/Staphylococcus group; Listeria.  
1 (sites)  
Glaser,P., Frangeul,L., Buchrieser,C., Rusnlok,C., Amend,A.,  
Baquero,F., Berche,P., Bloeker,H., Brandt,P., Chakraborty,T.,  
Charbit,A., Chetoui,F., Couve,E., de Daruvar,A., Deboux,P.,  
Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,  
Dusurget,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,  
Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,  
Jackson,D., Jones,L.M., Kaerst,U., Kref,J., Kuhn,M., Kunst,F.,  
Kurapkut,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,  
Nedjari,H., Nordstedt,G., Novella,S., de Pablo,B., Perez-Diaz,J.C.,  
Purcell,I., Remmel,B., Rose,M., Schlueter,T., Simoes,N.,  
Tierrez,A., Vazquez-Bolaj,J.A., Voss,H., Wehland,J. and Cossart,P.  
Comparative genomics of Listeria species  
Science 294 (5543), 849-852 (2001)  
21537279  
2 (bases 1 to 347050)  
Glaser,P., Frangeul,L. and Rusnlok,C.  
Direct Submission  
Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des  
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris  
Cedex 15, FRANCE  
E-mail: pglaser@pasteur.fr  
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.  
Location/Qualifiers  
1. 347050  
/organism="Listeria monocytogenes"  
/strain="EGD-e"  
/db\_xref="taxon:1639"  
complement(95. .841)  
/gene="lmo1688"  
complement(95. .841)  
/gene="lmo1688"  
/note="similar to glucose 1-dehydrogenase"  
/codon\_start=1  
/protein\_id="CAC99766.1"  
/db\_xref="GI:16411142"  
/translation="MNKVALYVGSRGIGREIALALAKEGDIAVNSRNKKAEEVQ  
QEIQLGRKCVIFRANVADVRELFKAVDEEGRDLIFINNAAGVLRPMLEES  
HMDWTMINAKALLFAGOEAKIMQHQSGKISISGSIYENNTYGVSKAVE  
SLTRIAVELAPGIAVNAVAGLIEFDALNHFPRRELLKDAVSKTPAGRMIEPNDL  
VNAVFLASEKADMRGOTILVDGRTLLV"  
complement(join(845. .1942,1948. .1953))  
/gene="lmo1689"  
complement(845. .1942)  
/note="similar to A/G-specific adenine glycosylase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC99767.1"  
/db\_xref="GI:16411143"  
/translation="MRKKRLTWDETKITAFQALVSYWEANKRVLPWRENTPEYRIN  
VSEIMLOOTKVDVTPYNNRPMTOPPTMESFNADLILKAMEGLGYSTRNLOTA  
KMOVADSGEPDILTIISIKGVPTAGAILISIAVNOAEPADGVNMYRIAYLE  
IGEDIMKASTRIPEEVIKOLIDKNPAPFNGIMLEIALVCTPTKPMCMCLDPLPFC  
EAKMGVETNPVKIKVKMKTKELITIVISEDKIAIERKPEGLIANMWFPTIE

gene  
complement(850. .855)  
/gene="lmo1688"  
complement(850. .855)  
/gene="lmo1688"  
complement(1948. .1953)  
/gene="lmo1689"  
join(2129. .3123,3130. .3147)  
/gene="lmo1690"  
2129. .2134  
/gene="lmo1690"  
2143. .3123  
/gene="lmo1690"  
/note="similar to hypothetical proteins"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC99768.1"  
/db\_xref="GI:16411144"  
/translation="MDGTTHVVMGIALGALATVDPVVAGSSQAIGIMTATIGSQIP  
DIDVTLKLNADYIRNHRGITHSLPMLAIPMLISSILYLPATFTHLLMTEIA  
VGLHLEVDENAGYQAVRPEKTEWVAFGEIWTPEWIRGSHVVAIAAMLGSPVLT  
EVLIVYIILAYVARRVQRMKHAVONIPSEETIISTHIFQMRVAVYTKHYY  
VGRAKRNTSIEKEDRLPYDPNEIIRSAKKRNIAAFISFKVYMWREELDGTIV  
TFTDLKRNNGHYPRVAVVKLDDLVSSYTGWIFSTEKLKLAIVSI"  
complement(join(3130. .3147,3156. .3617,3625. .3630))  
/gene="lmo1691"  
3130. .3147  
/gene="lmo1690"  
complement(3156. .3617)  
/gene="lmo1691"  
/note="similar to deoxyuridine triphosphate  
nucleotidohydrolases"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC99769.1"  
/db\_xref="GI:16411145"  
/translation="MKVGFVNVNENSRKPEQTISLPTRGDKSGAGYDFPSENYTI  
VPEKHFMTDLSKYNQOEDEVNITYRSSIGIKKGLLNGTGIIISYSSPNDGNG  
IGAIKPNSENVSTIAGRVAGVQKTLVADTDIVANESVGGVSGISR"  
complement(3625. .3630)  
/gene="lmo1691"  
complement(3629. .3862)  
/gene="lmo1692"  
complement(3629. .3862)  
/gene="lmo1692"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC99770.1"  
/db\_xref="GI:16411146"  
/translation="MTYTLDELTAELMWNPPLDKEQALMWIEMLMSDFESSYAKGY  
PYRGPEYADYRQOIERRGSLFHOYERDPKK"  
complement(join(3864. .4673,4680. .4685))  
/gene="lmo1693"  
complement(3864. .4673)  
/gene="lmo1693"  
/note="similar to hypothetical proteins"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC99771.1"  
/db\_xref="GI:16411147"  
/translation="MKITSISVQKNKERYNIFIDEKYNFSDVEYLARYOLMKGVL  
TEADIEETKQADMVRKGLNKAIKSHRVSREKRLDYRKQEMEPFADILKIKKAN  
MDYINDPEFALYTKQITKIGPRTIEREYVKGUTREITISRYEESDQAQLENA  
TKQAIKTRNNKSAKMLDQKLTIDLOKGFSTELAKAATENATSELSADAEALIQ  
KOIEKTRNNKRNKPSIAKOKTITSLMOKGFSYDITIOSLYTENISFEEREE"  
complement(3872. .3877)  
/gene="lmo1692"  
complement(3872. .3877)  
/gene="lmo1692"  
complement(4680. .4685)

```

gene      /gene="lmo1693"
          join(4757..5673,5681..5708)
          /gene="lmo1694"
          4757..4762
          /gene="lmo1694"
          4771..5673
          /gene="lmo1694"
          /note="similar to CDP-abequose synthase"
          /codon_start=1
          /transl_table=11
          /protein_id="CAC99772.1"
          /db_xref="GI:16411148"
          /translation="MHILFGATGFTGDHLVHELEKSDHELYILTRQIKKANPANHVI
          EMLNEDKLPNLEDPVDCINLAGAGLMDERWTYDRKVIYNSRIEASALLSYKMK
          KSKPKLWLNAGAIYSSKSTIYLTDEENTYADNLTQVEMDEKTSASADLIRV
          VYARFGLVINGSGSPVEKLFQYTGRCNGRQWYSMIHVDVVAILETFDHEO
          INGVAFTAPHPVOEKKEFKRGLGKMKHPRYPVPKIKIKFILGERAMTILDSQARF
          EKLSNHFREFETLQEAIDLDD"
          5681..5708
          /gene="lmo1694"
          complement(join(5682..8291,8239..8304))
          /gene="lmo1695"
          complement(5682..5709)
          /gene="lmo1695"
          complement(5694..8291)
          /gene="lmo1695"
          /note="similar to putative membrane proteins"
          /codon_start=1
          /transl_table=11
          /protein_id="CAC99773.1"
          /db_xref="GI:16411149"
          /translation="MKELMQAYAFQKNSYVVKVTFITFVAWVFEIINATIDY
          PSLKENTLSQSEPOIFIMFIVGLIATVPMIYDYIVLPGKFSPIVIAAGTINT
          FTNKGFGVGLVSGISRASFSYKNSHKEILALISLAEIVSGLSITCVSLATLIP
          GFADHPVNWPMILAGGLYEPILFTITKSKSLDPIKRELTLIILASILEMCFAR
          GCPALITGMEPYDIPKVPYIYASVIGTASWPGVGTFFDYMIIGLSQISQAF
          LALAMLPFRITFTIPVGLFVQKAGKNDLEGLPLFLQKVAHRELIVFY
          GSGLLILSSAVPNAIYHVPFLYKIMPNFLFTSOITTVAFGLLGARIECTKYK
          AYITIVIVGCAIFNTLAVRSMKQALFGLVLCFLFARNFVREKLYVTSKYIID
          SIITVCLAGIVIGVNSPNIKHSKEIPDYLRILASEHMLMGVGFVLAIVYIIV
          IYLTSTKELGSPPEAVVVRHKLAKNGENSHFTFLDKLLEFMAEGEVLESYRTIA
          DKVIMGEPTGNMCKMEALTEVMMNADRPGRVPEVYVRGMPTYLIDHGPDKIG
          BEGVVDQNTFMSKKKKKKGERALNKKLEREGTEFTIIPERNHDTWTTLRAVSDMLD
          GREKGFSLGFDYTYLEQAPIALAKNGEFTVGSAMPSTYDGMTSIDLRYSKEA
          PSGIMDLFLINLFEKAKEDGFOTNAGMAPLANWESYAEFGEGLAGLVYRSGFY
          GFGLRNFKSKYVTEWQKFVAFRRRSSIAFTMLQMLTVGKKRPLANSOVVLDPELE
          EETKKPDSE"
          complement(8239..8304)
          /gene="lmo1695"
          complement(8318..9007)
          /gene="lmo1696"
          complement(8318..8992)
          /gene="lmo1696"
          /note="similar to unknown proteins"
          /codon_start=1
          /transl_table=11
          /protein_id="CAC99774.1"
          /db_xref="GI:16411150"
          /translation="MOYOKCRFVLLPALYLGLISALOGFNADLINTFANSCL
          FLATLITLNARLKNWIDTWEVCVPIIYIILLHVAIAGDPVNSGTAGTTLIYIF
          MINIPPTTIENTFOQLPTPTPTIIOIFGNVLLCPSFFMAYRITSTAGTTLIYIF
          LITSGIILLDPAQTMITGFESISLPPKRSYDIDIIILNTLSGLILLALVAPSVRK

```

Query Match 1.8%; Score 21; DB 1; Length 34705;  
 Best Local Similarity 100.0%; Pred. No. 6.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 147 atccgtaaggtatcatcca 167  
 |||||  
 Db 224218 ATCCGTAAGGTATCATCCA 224238

RESULT 14

```

AF182790/c 573 bp DNA linear PLN 06-DEC-1999
LOCUS
DEFINITION
Pueraria phaseoloides cytochrome oxidase subunit 2 (cox2) nuclear
pseudogene, exon 2, partial sequence.
ACCESSION
AF182790
VERSION
AF182790.1 GI:6525126
KEYWORDS
SOURCE
ORGANISM
Pueraria phaseoloides.
Pueraria phaseoloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Pueraria.
REFERENCE
1 (bases 1 to 573)
Adams,K.L., Song,K., Roessler,P.G., Nugent,J.M., Doyle,J.L.,
Doyle,J.J. and Palmer,J.D.
Intracellular gene transfer in action: dual transcription and
multiple silencings of nuclear and mitochondrial cox2 genes in
legumes
Proc. Natl. Acad. Sci. U.S.A. 96 (24), 13863-13868 (1999)
JOURNAL
MEDLINE
20040642
REFERENCE
2 (bases 1 to 573)
Song,K.
Direct Submission
Submitted (02-SEP-1999) Biology, Indiana University, Jordan Hall
142, 1001 E. Third St., Bloomington, IN 47405, USA
FEATURES
source
1..573
/organism="Pueraria phaseoloides"
/db_xref="taxon:109224"
<1..>573
/gene="cox2"
/pseudo
/number=2
<1..>573
/gene="cox2"
/note="cytochrome oxidase subunit 2"
/pseudo
BASE COUNT
166 a 107 c 113 g 187 t
ORIGIN
Query Match 1.8%; Score 20; DB 8; Length 573;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 543 tctgaatgataaagaaag 562
|||||
Db 137 tctgaatgataaagaaag 118
RESULT 15
MMU272046 636 bp mRNA linear ROD 04-APR-2000
MMU272046
Mus musculus mRNA for calcium channel gamma 5 subunit (CACNG5
gene).
ACCESSION
AJ272046
VERSION
AJ272046.1 GI:7452997
KEYWORDS
CACNG5 gene; calcium channel gamma 5 subunit.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Klugbauer,N., Dal,S., Specht,V., Lacinova,L., Marais,E., Bohn,G.
and Hofmann,E.
A family of gamma-like calcium channel subunits
FEBS Lett. 470 (2), 189-197 (2000)
JOURNAL
MEDLINE
20200313
REFERENCE
2 (bases 1 to 636)
Klugbauer,N.
Direct Submission
Submitted (08-FEB-2000) Klugbauer N., Institut fuer Pharmakologie

```

und Toxikologie, Technische Universitaet Muenchen, Biedersteiner  
Str. 29, 81827 Muenchen, GERMANY

FEATURES  
source

location/Qualifiers  
1..636  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="kidney"

gene

1..636  
/gene="CACNG5"

CDS

1..636  
/gene="CACNG5"  
/codon\_start=1  
/product="calcium channel gamma 5 subunit"  
/protein\_id="CAB86387.1"  
/db\_xref="GI:7452998"  
/translation="MTAIGAQAHLGLKRPFRSFESFIRTLIIYCTALAVLSYS  
ICDGHMLVEDHLEFCTIGNHSEPHCLRDLSOAHMPGLAVGMGLARSVAMAVY  
AAIFGLMLIVSOVCEDVRSRRKVAIGSYLLIVAFILSSGGLPFIILIKKQINLGF  
TLMFCEFTASFLFFLNASGLHINSILTPWDPPTLAVLRKRGIDGTSIL"

BASE COUNT 107 a 204 c 163 g 162 t  
ORIGIN

Query Match 1.8% Score 20; DB 10; Length 636;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 938 tgggcttcacccctgatgttc 957  
|||||  
Db 473 TGGGCTTCACCCCTGATGTTC 492

Search completed: May 7, 2002, 18:10:17  
Job time: 10585 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 16:47:58 ; Search time 248.07 seconds

(without alignments)  
7869.278 Million cell updates/sec

Title: US-09-579-383-1

Perfect score: 1137

Sequence: 1 atgaatttcacgtaaata.....tcgtgaagccctgcactaa 1137

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_032802:\*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1137	100.0	1137	22	AAC89668
2	22	1.9	2122	22	AAF60839
3	20	1.8	746	24	ABL01346
4	19	1.7	882	24	AAK61604
5	19	1.7	1282	21	AAC50729
6	19	1.7	1289	21	AAC40124
7	18	1.6	453	18	AAV74939
8	18	1.6	581	22	ABA63976
9	18	1.6	581	22	ABA31149

c	10	18	1.6	581	22	AAK12475	Human brain expres
c	11	18	1.6	581	22	AAK38184	Human bone marrow
c	12	18	1.6	581	22	AAI18978	Probe #8911 for ge
c	13	18	1.6	581	22	AAI44114	Probe #12800 used
c	14	18	1.6	694	21	AAI12222	Aspergillus oryzae
c	15	18	1.6	1396	22	AAI58814	Human cDNA encodin
c	16	18	1.6	1684	23	ABL09429	Drosophila melanog
c	17	18	1.6	1858	22	AAF81671	Human membrane ass
c	18	18	1.6	1911	23	ABL02868	Drosophila melanog
c	19	18	1.6	1961	22	ABA61573	Human breast cell
c	20	18	1.6	1961	22	ABA56719	Human breast liver
c	21	18	1.6	1961	22	AAI14984	Probe #4917 for ge
c	22	18	1.6	1961	22	AAI36330	Probe #5016 used t
c	23	18	1.6	1986	22	ABA61618	Human breast liver
c	24	18	1.6	1986	22	ABA56714	Human breast cell
c	25	18	1.6	1986	22	AAK26329	Probe #4795 for ge
c	26	18	1.6	1986	22	AAK04844	Human brain expres
c	27	18	1.6	1986	22	AAI14980	Probe #4913 for ge
c	28	18	1.6	1986	22	AAI36326	Probe #5012 used t
c	29	18	1.6	1986	22	AAI04747	Probe #4738 used t
c	30	18	1.6	2081	21	AAV78399	Human secreted pro
c	31	18	1.6	2334	22	ABA51285	Human breast cell
c	32	18	1.6	2334	22	ABA69293	Human foetal liver
c	33	18	1.6	2334	22	AAI24172	Probe #14105 for g
c	34	18	1.6	2334	22	AAI49459	Probe #18145 used
c	35	18	1.6	2351	22	ABA51280	Human breast cell
c	36	18	1.6	2351	22	ABA69288	Human foetal liver
c	37	18	1.6	2351	22	ABA36214	Probe #14680 for g
c	38	18	1.6	2351	22	AAK17575	Human brain expres
c	39	18	1.6	2351	22	AAI24168	Probe #14101 for g
c	40	18	1.6	2351	22	AAI49455	Probe #18141 used
c	41	18	1.6	2351	22	AAI09732	Probe #9723 used t
c	42	18	1.6	2723	23	ABL26896	Drosophila melanog
c	43	18	1.6	3164	21	AAZ58979	Human cytoskeletal
c	44	18	1.6	3263	23	AA584443	DNA encoding novel
c	45	18	1.6	3416	13	AAQ21004	hek receptor prote
c	46	18	1.6	4258	18	ABL18134	Drosophila melanog
c	47	18	1.6	4419	22	AAD06838	Human STRAP-3 cDNA
c	48	18	1.6	4429	22	AAD07073	Human six transmem
c	49	18	1.6	6614	20	AAZ69905	6.6 kb transcript
c	50	18	1.6	7412	23	ABL01984	Drosophila melanog
c	51	18	1.6	11131	23	ABL09428	Drosophila melanog
c	52	18	1.6	12438	23	ABL20810	Drosophila melanog
c	53	18	1.6	16092	23	ABL05124	Drosophila melanog
c	54	18	1.6	31952	22	AAK89370	Human digestive sy
c	55	18	1.6	32250	22	AAI06975	Human reproductive
c	56	18	1.6	38342	22	AA546745	Tumour suppressor
c	57	18	1.5	188	21	AA41572	Human secreted exp
c	58	17	1.5	269	21	AAK16548	Human secreted pro
c	59	17	1.5	300	20	AAK98380	Human cancer cell
c	60	17	1.5	314	22	AAH99757	Human protein enco
c	61	17	1.5	351	22	ABA16776	Human nervous syst
c	62	17	1.5	479	21	AAC40951	zebra mays DNA fragm
c	63	17	1.5	615	18	AAI84082	DNA encoding a Sta
c	64	17	1.5	696	22	AAF56442	Neisseria meningit
c	65	17	1.5	696	22	AAF56443	Neisseria meningit
c	66	17	1.5	696	22	AAF56444	Neisseria meningit
c	67	17	1.5	696	22	AAF56445	Neisseria meningit
c	68	17	1.5	696	22	AAF56446	Neisseria meningit
c	69	17	1.5	696	22	AAF56447	Neisseria meningit
c	70	17	1.5	696	22	AAF56448	Neisseria meningit
c	71	17	1.5	696	22	AAF56449	Neisseria meningit
c	72	17	1.5	696	22	AAF56450	Neisseria meningit
c	73	17	1.5	696	22	AAF56451	Neisseria meningit
c	74	17	1.5	696	22	AAF56452	Neisseria meningit
c	75	17	1.5	696	22	AAF56453	Neisseria meningit
c	76	17	1.5	696	22	AAF56454	Neisseria meningit
c	77	17	1.5	714	21	AAK05483	Streptococcus pneu
c	78	17	1.5	747	21	AAZ91827	Human neuroblastom
c	79	17	1.5	809	22	AAI22553	Human breast cance
c	80	17	1.5	849	22	AAH64605	E. coli growth and
c	81	17	1.5	880	23	ABL16495	Drosophila melanog
c	82	17	1.5				

c 83	17	1.5	924	22	AAK57408	Human immune/haema
c 84	17	1.5	942	22	AAH74584	Synthetic nucleoti
c 85	17	1.5	957	14	AAQ51347	c424 gene, Asperg
c 86	17	1.5	969	20	AAAX19904	Streptococcus pneu
c 87	17	1.5	1016	23	ABLO7413	Drosophila melanog
c 88	17	1.5	1032	22	AAH29058	Drosophila melanog
c 89	17	1.5	1035	21	AAK34605	Arabidopsis thalia
c 90	17	1.5	1114	22	ABAA9026	Human breast cell
c 91	17	1.5	1114	22	ABA34028	Probe #12494 for g
c 92	17	1.5	1114	22	AAK15385	Human brain expres
c 93	17	1.5	1114	22	AAK41110	Human bone marrow
c 94	17	1.5	1114	22	AAK121874	Probe #11807 for g
c 95	17	1.5	1114	22	AAI07562	Probe #7553 used t
c 96	17	1.5	1177	21	AAK41665	Arabidopsis thalia
c 97	17	1.5	1226	22	AAK91718	Human polynucleoti
c 98	17	1.5	1370	19	AAZ96377	S. pneumoniae deri
c 99	17	1.5	1490	22	AAK1021	DNA encoding novel
c 100	17	1.5	1490	22	ABAA34927	Human breast cell
c 101	17	1.5	1490	22	ABAA24160	Probe #2626 for ge
c 102	17	1.5	1490	22	AAK02675	Human bone marrow
c 103	17	1.5	1490	22	AAK28113	Probe #2623 for ge
c 104	17	1.5	1490	22	AAI12690	Probe #2590 used t
c 105	17	1.5	1490	22	AAI02599	DNA encoding novel
c 106	17	1.5	1600	23	AA572685	WO 99/07855 Seqid
c 107	17	1.5	1627	20	AAK222900	Human polynucleoti
c 108	17	1.5	1630	22	AAI59400	WO 99/07855 Seqid
c 109	17	1.5	1662	20	AAK22899	Human polynucleoti
c 110	17	1.5	1668	22	AAI61186	Human polynucleoti
c 111	17	1.5	1668	22	AAI61187	Human polynucleoti
c 112	17	1.5	1666	23	ABLO8451	Drosophila melanog
c 113	17	1.5	1666	23	ABLO84426	DNA encoding novel
c 114	17	1.5	1759	22	AAI59401	Human polynucleoti
c 115	17	1.5	1780	21	AAK41739	Arabidopsis thalia
c 116	17	1.5	1988	21	AAK03309	Human brain expres
c 117	17	1.5	2193	21	AAK68671	Rat R35 CDNA. Rat
c 118	17	1.5	2565	22	AAK91715	Human polynucleoti
c 119	17	1.5	2568	21	AAK21637	Human breast and o
c 120	17	1.5	2686	11	AAO10163	Cyclomatidoxetin
c 121	17	1.5	2687	11	AAO01792	Sequence encoding
c 122	17	1.5	2863	11	AAO05631	Lambda gtl0ch2 enc
c 123	17	1.5	2976	21	AAA26914	Essential1 Staphylo
c 124	17	1.5	2976	22	AA508065	Staphylococcus aur
c 125	17	1.5	2976	22	AAH17705	Human cDNA sequenc
c 126	17	1.5	2976	22	AAK91596	Staphylococcus aur
c 127	17	1.5	3194	23	ABLO7412	Drosophila melanog
c 128	17	1.5	3217	23	ABLI6494	Drosophila melanog
c 129	17	1.5	3274	23	ABLI9018	Drosophila melanog
c 130	17	1.5	3316	11	AAO05632	Lambda gtl0ch101 e
c 131	17	1.5	3516	22	AAK85421	Human immune/haema
c 132	17	1.5	3527	23	ABLO9309	Drosophila melanog
c 133	17	1.5	3688	11	AAO05630	Lambda gtl0ch201 e
c 134	17	1.5	3703	13	AAO20200	Gene encoding C-te
c 135	17	1.5	3712	23	ABLO7600	Drosophila melanog
c 136	17	1.5	3746	10	AAK90693	Bovine peptidyl-gl
c 137	17	1.5	3844	13	AAQ20199	Gene encoding C-te
c 138	17	1.5	3904	13	AAQ20198	Gene encoding C-te
c 139	17	1.5	3925	12	AAQ11041	Encodes equine C-t
c 140	17	1.5	4045	13	AAQ20197	Gene encoding C-te
c 141	17	1.5	4370	23	ABLO8450	Drosophila melanog
c 142	17	1.5	4563	24	AA594875	Human DNA sequence
c 143	17	1.5	4676	23	AA584690	DNA encoding novel
c 144	17	1.5	4732	22	AAK70267	Human immune/haema
c 145	17	1.5	5066	19	AAV52212	Streptococcus pneu
c 146	17	1.5	5137	18	AAV74528	Staphylococcus aur
c 147	17	1.5	5274	22	AAK6782	Human immune/haema
c 148	17	1.5	5420	23	ABLI27968	Drosophila melanog
c 149	17	1.5	5836	23	ABLO2526	Drosophila melanog
c 150	17	1.5	6231	23	ABLO9308	Drosophila melanog
c 151	17	1.5	6558	22	ABA21372	Human nervous syst
c 152	17	1.5	6558	22	AAI37138	Human musculoskele
c 153	17	1.5	6558	22	AAK86101	Human immune/haema
c 154	17	1.5	7953	18	AAV74463	Staphylococcus aur
c 155	17	1.5	8367	21	AAK81483	N. meningitidis pa
c 156	17	1.5	11093	23	ABLO2506	Drosophila melanog
c 157	17	1.5	24200	23	ABLI21182	Drosophila melanog
c 158	17	1.5	24206	23	ABLI21180	Drosophila melanog
c 159	17	1.5	24235	23	ABLO5026	Drosophila melanog
c 160	17	1.5	30889	23	ABLI9852	Drosophila melanog
c 161	17	1.5	31563	23	ABLO2738	Drosophila melanog
c 162	17	1.5	34094	20	AAZ30165	Complete nucleotid
c 163	17	1.5	38189	23	ABLI9856	Drosophila melanog
c 164	17	1.5	39814	23	ABLI9854	Drosophila melanog
c 165	17	1.5	40352	19	AAV02032	MAGE-B cluster DNA
c 166	17	1.5	48974	20	AAK55300	Mouse Presentin-1
c 167	17	1.5	349980	21	AAK21544	Neisseria meningit
c 168	17	1.5	349980	21	AAK21607	Neisseria meningit
c 169	17	1.5	1437668	21	AAK81490	N. meningitidis B
c 170	17	1.5	164976	19	AAV21209	Methanococcus jan
c 171	17	1.4	163	22	AAI12297	5' UTR of yeast YA
c 172	17	1.4	178	22	AAI13911	Human secreted pro
c 173	16	1.4	187	21	AAK31265	Human bone marrow
c 174	16	1.4	193	21	AAK51355	Human secreted pro
c 175	16	1.4	230	21	AAK04981	Human secreted pro
c 176	16	1.4	231	24	AA562677	Human MSR2 gene ex
c 177	16	1.4	248	16	AAQ93948	Human MSR2 gene ex
c 178	16	1.4	248	16	ABLO1821	Human MSR2 (BMSH2)
c 179	16	1.4	251	16	AAK23682	Human gene signatu
c 180	16	1.4	274	22	AAI12196	Human breast cance
c 181	16	1.4	296	21	AAK67567	Eucalyptus grandis
c 182	16	1.4	300	20	AAI14867	Human gene expres
c 183	16	1.4	315	22	ABA70985	Human foetal liver
c 184	16	1.4	315	22	ABA37403	Probe #15869 for g
c 185	16	1.4	315	22	AAK19261	Human brain expres
c 186	16	1.4	315	22	AAK45231	Human bone marrow
c 187	16	1.4	315	22	AAI25202	Probe #15135 for g
c 188	16	1.4	320	22	AAI51182	Probe #1968 used
c 189	16	1.4	320	22	AAK57185	Human immune/haema
c 190	16	1.4	321	16	AAK93398	Human MSR2 gene PC
c 191	16	1.4	339	22	AA530328	DNA encoding novel
c 192	16	1.4	339	22	AAI02009	Human reproductive
c 193	16	1.4	344	22	AAK63302	Human immune/haema
c 194	16	1.4	346	21	AAK67566	Eucalyptus grandis
c 195	16	1.4	359	21	AAK26658	Human secreted pro
c 196	16	1.4	370	21	AAK98933	Human pancreatic c
c 197	16	1.4	375	21	AAK40202	H. pylori 26 kDa p
c 198	16	1.4	377	22	AAK63027	Human immune/haema
c 199	16	1.4	385	22	AAI21079	Human breast cance
c 200	16	1.4	391	21	AAK26965	Human secreted pro
c 201	16	1.4	395	22	AAK75385	Human immune/haema
c 202	16	1.4	400	22	AAI82249	Human polynucleoti
c 203	16	1.4	401	22	AAK96258	Human neutrophilin g
c 204	16	1.4	401	22	AAK97751	Human neutrophilin g
c 205	16	1.4	402	22	AAK96724	Human neutrophilin g
c 206	16	1.4	402	22	AAK98217	Human neutrophilin g
c 207	16	1.4	402	22	AAK59471	Human immune/haema
c 208	16	1.4	406	21	AAK67561	Eucalyptus grandis
c 209	16	1.4	426	22	AAK74756	Human immune/haema
c 210	16	1.4	426	22	AAK74757	Human immune/haema
c 211	16	1.4	431	22	AAK59220	Human immune/haema
c 212	16	1.4	440	23	AAK90103	DNA encoding novel
c 213	16	1.4	445	22	AAK65711	Novel human polynu
c 214	16	1.4	451	22	AAK48952	Human bone marrow
c 215	16	1.4	455	22	AAK58389	Human foetal liver
c 216	16	1.4	455	22	ABA27498	Probe #5964 for ge
c 217	16	1.4	455	22	AAK06494	Human bone marrow
c 218	16	1.4	455	22	AAK32175	Human brain expres
c 219	16	1.4	455	22	AAI15985	Probe #5918 for ge
c 220	16	1.4	455	22	AAI38024	Probe #6710 used t
c 221	16	1.4	456	22	AAK60779	Human immune/haema
c 222	16	1.4	475	22	ABA51960	Human foetal liver
c 223	16	1.4	475	22	ABA51961	Human foetal liver
c 224	16	1.4	475	22	ABA21778	Probe #244 for gen
c 225	16	1.4	475	22	ABA21779	Probe #245 for gen
c 226	16	1.4	475	22	AAK00246	Human brain expres
c 227	16	1.4	475	22	AAK00247	Human brain expres
c 228	16	1.4	475	22	AAK25687	Human bone marrow



229	16	1.4	475	22	AAK25688	Human bone marrow
230	16	1.4	475	22	AAI10315	Probe #248 for gen
231	16	1.4	475	22	AAI10316	Probe #248 for gen
232	16	1.4	475	22	AAI31568	Probe #254 used to
233	16	1.4	475	22	AAI31569	Probe #254 used to
234	16	1.4	475	22	AAI00252	Probe #242 used to
235	16	1.4	475	22	AAI00251	Probe #243 used to
236	16	1.4	480	20	AAV86552	EST clone AW170.
237	16	1.4	500	23	AA576003	DNA encoding novel
238	16	1.4	538	22	AAI02558	Human reproductive
239	16	1.4	543	22	AAI34807	Human musculoskele
240	16	1.4	546	22	AAH35057	Human colon cancer
241	16	1.4	549	22	AAK56724	Human immune/haema
242	16	1.4	553	22	AAH10327	Human cDNA clone (
243	16	1.4	568	22	ABA06963	Human pancreatic c
244	16	1.4	568	22	AAK88162	Human digestive sy
245	16	1.4	582	22	AAK38542	Human bone marrow
246	16	1.4	589	22	AA542810	Human G Protein-Co
247	16	1.4	589	22	AAK36030	Human bone marrow
248	16	1.4	593	22	ABA61912	Human foetal liver
249	16	1.4	593	22	AAK10222	Human brain expres
250	16	1.4	593	22	AAK36121	Human bone marrow
251	16	1.4	593	22	AAI41835	Probe #10521 used
252	16	1.4	618	18	AAV24658	H. pylori ORF 05cp
253	16	1.4	624	21	AAZ93841	RAB6c coding sequ
254	16	1.4	715	21	AAK50979	Arabidopsis thalia
255	16	1.4	718	21	AAZ93835	RAB6 coding sequen
256	16	1.4	721	21	AAK40852	Arabidopsis thalia
257	16	1.4	723	22	AA563224	DNA encoding human
258	16	1.4	723	22	AAK91834	Human cDNA 5'-end
259	16	1.4	733	22	AA523835	E. coli DNA for ce
260	16	1.4	739	21	AAZ93836	RAB6c coding sequ
261	16	1.4	740	21	AAZ93837	RAB6 coding sequen
262	16	1.4	747	23	AA575716	DNA encoding novel
263	16	1.4	766	20	AAK39666	Renal cancer assoc
264	16	1.4	791	23	AA572736	DNA encoding novel
265	16	1.4	793	22	AAK93184	Human cDNA clone r
266	16	1.4	798	23	AA566629	DNA encoding novel
267	16	1.4	798	23	AA576013	DNA encoding novel
268	16	1.4	798	23	AA581257	DNA encoding novel
269	16	1.4	828	20	AAK61652	B. burgdorferi ant
270	16	1.4	833	22	AA533006	DNA encoding novel
271	16	1.4	833	22	AA533007	DNA encoding novel
272	16	1.4	849	18	AAV06127	Viral infection ge
273	16	1.4	860	20	AAK39911	Gastric cancer ass
274	16	1.4	873	20	AAK61651	B. burgdorferi ant
275	16	1.4	875	22	AA57746	Corynebacterium gl
276	16	1.4	885	22	AA573262	S. epidermidis ope
277	16	1.4	888	20	AAK57388	Rat U3 gene trap d
278	16	1.4	921	22	AAH24227	Human oxidoreducta
279	16	1.4	956	22	AAZ32221	Sp5 mutant clone w
280	16	1.4	957	22	AAH31775	Human olfactory re
281	16	1.4	960	22	AA542276	Human cDNA encodin
282	16	1.4	965	8	AAH70775	Sequence encoding
283	16	1.4	966	13	AAQ29960	Sugar beet chitina
284	16	1.4	996	22	AA529636	DNA encoding P. ae
285	16	1.4	1001	20	AAK03300	Partial cDNA encod
286	16	1.4	1065	22	AA571894	Corynebacterium gl
287	16	1.4	1076	22	AA579935	Human apoptosis-re
288	16	1.4	1107	22	AAH26036	DNA encoding spina
289	16	1.4	1110	18	AAH75411	Staphylococcus aur
290	16	1.4	1158	23	AA570851	DNA encoding novel
291	16	1.4	1172	23	AA544981	CDNA encoding nove
292	16	1.4	1197	22	AAI66530	Human pterio-molyb
293	16	1.4	1239	22	AAK68963	Human immune/haema
294	16	1.4	1247	13	AAQ21978	Hybrid mini-vitel
295	16	1.4	1256	16	AAQ91638	Zebrafish sonic he
296	16	1.4	1256	16	AAK25621	Zebrafish sonic he
297	16	1.4	1256	20	AAK25102	Zebrafish sonic he
298	16	1.4	1256	20	AAK07275	Zebrafish sonic he
299	16	1.4	1256	20	AAK16186	Zebrafish Shh hedg
300	16	1.4	1256	20	AAK16186	Zebrafish Sonic he
301	16	1.4	1256	21	AA50450	Zebrafish Sonic he
302	16	1.4	1256	21	AAK27880	Zebrafish Sonic he
303	16	1.4	1256	21	AAK30278	Partial zebrafish
304	16	1.4	1256	21	AAK25261	Zebrafish Sonic he
305	16	1.4	1256	22	AAI66775	Zebrafish sonic he
306	16	1.4	1256	22	AAH76111	Zebrafish sonic he
307	16	1.4	1256	22	AAH28450	Nucleotide sequenc
308	16	1.4	1256	22	AAAD09033	Zebrafish sonic he
309	16	1.4	1256	22	AAAD10150	Nucleotide sequenc
310	16	1.4	1256	22	AA572017	Zebrafish Sonic he
311	16	1.4	1256	22	AA572017	Zebrafish Sonic he
312	16	1.4	1256	24	AAAD23803	Hybrid mini-vitel
313	16	1.4	1260	13	AAQ21976	Hybrid mini-vitel
314	16	1.4	1290	13	AAQ21977	Hybrid mini-vitel
315	16	1.4	1290	13	AAQ21980	Hybrid mini-vitel
316	16	1.4	1290	13	AAQ21981	Hybrid mini-vitel
317	16	1.4	1317	22	AA575387	S. epidermidis ope
318	16	1.4	1329	21	AA571179	Human ORFX ORF2734
319	16	1.4	1334	21	AA515805	Human prostate can
320	16	1.4	1335	22	AAH52115	Human AFP protein
321	16	1.4	1349	22	AAI99526	Human polynucleoti
322	16	1.4	1374	23	AA573577	DNA encoding novel
323	16	1.4	1383	22	AAH66506	C. glutamicum codin
324	16	1.4	1393	7	AAH60225	Sequence of autono
325	16	1.4	1405	14	AAQ46125	PE2 receptor (EP3
326	16	1.4	1432	22	AAK94247	Human full-length
327	16	1.4	1444	22	AAH44073	Oryza sativa perox
328	16	1.4	1455	21	AAAI5005	CDNA encoding a hu
329	16	1.4	1460	21	AAZ93840	RAB6c coding sequ
330	16	1.4	1461	23	ABL18553	Drosophila melanog
331	16	1.4	1506	23	ABL05251	H. pylori secreted
332	16	1.4	1509	18	AAV24861	H. pylori secreted
333	16	1.4	1509	18	AAV24959	Human WAVE3 CDNA
334	16	1.4	1509	22	AA572913	DNA encoding novel
335	16	1.4	1523	22	AA522913	DNA encoding novel
336	16	1.4	1525	22	AA522912	DNA encoding novel
337	16	1.4	1583	24	ABA03464	Human DNA mismatch
338	16	1.4	1590	23	AA593665	Arabidopsis thalia
339	16	1.4	1591	21	AAK33449	DNA encoding novel
340	16	1.4	1620	23	AA570699	DNA encoding novel
341	16	1.4	1620	23	AA590727	DNA encoding novel
342	16	1.4	1622	23	AA587607	Human polynucleoti
343	16	1.4	1625	22	AAI99513	Human polynucleoti
344	16	1.4	1630	19	AAK14097	H. pylori GHP0 137
345	16	1.4	1638	21	AAK51437	Arabidopsis thalia
346	16	1.4	1644	22	AA522750	Human cDNA encodin
347	16	1.4	1660	21	AAZ65008	Membrane-bound pro
348	16	1.4	1660	22	AA545998	Human DNA encoding
349	16	1.4	1660	22	AA592073	Human PRO1115 CDNA
350	16	1.4	1660	22	AA544154	Human PRO1115 (UNQ
351	16	1.4	1710	22	AA561026	P. putida KT2440-a
352	16	1.4	1713	17	AA573689	Atyl Beta-N-acetyl
353	16	1.4	1713	21	AAZ38242	Vibrio furnissii e
354	16	1.4	1724	21	AA515819	Human prostate can
355	16	1.4	1726	22	AA522514	Human cDNA encodin
356	16	1.4	1755	22	AA589200	Novel human membra
357	16	1.4	1758	22	AA589199	Novel human membra
358	16	1.4	1764	23	AA5659058	DNA encoding novel
359	16	1.4	1770	21	AA561279	Arabidopsis thalia
360	16	1.4	1773	23	AA590107	DNA encoding novel
361	16	1.4	1794	23	AA570895	DNA encoding novel
362	16	1.4	1794	23	AA591006	DNA encoding novel
363	16	1.4	1887	22	AAH18090	Human cDNA sequenc
364	16	1.4	1893	24	AA562571	CDNA sequence #358
365	16	1.4	1967	23	ABL07527	Drosophila melanog
366	16	1.4	1968	22	AA521325	Human cDNA sequenc
367	16	1.4	2065	22	AA572219	Human neovasculari
368	16	1.4	2095	23	ABL13403	Drosophila melanog
369	16	1.4	2107	14	AAQ46124	PE2 receptor (EP3
370	16	1.4	2158	12	AAQ14624	Plasmid pPATDP in
371	16	1.4	2158	17	AAQ14372	Plasmid pPATDP (A
372	16	1.4	2158	20	AAZ32229	Rat dunce-like pho
373	16	1.4	2158	20	AAK88164	Plasmid pPATDP 2.
374	16	1.4	2161	20	AAK57464	Rat U3 gene trap d

c 375	16	1.4	2173	19	AAV31987	448	16	1.4	3452	22	AAK75977	Human immune/haema
c 376	16	1.4	2181	21	AAZ99576	449	16	1.4	3504	22	AAK52083	Human polynucleoti
c 377	16	1.4	2182	24	AAZ59595	c 450	16	1.4	3506	22	ABL05250	Drosophila melanog
c 378	16	1.4	2192	23	AAST1438	451	16	1.4	3592	23	AA584689	DNA encoding novel
c 379	16	1.4	2213	24	AA599922	c 452	16	1.4	3720	22	AAK53057	Human polynucleoti
c 380	16	1.4	2223	13	AAQ29187	453	16	1.4	3734	22	AA572406	DNA encoding novel
c 381	16	1.4	2223	16	AAQ80908	454	16	1.4	3763	19	AAV58194	Human myosin I-cha
c 382	16	1.4	2255	21	AAQ6514	455	16	1.4	3763	19	AA569872	DNA encoding novel
c 383	16	1.4	2274	17	AAQ75959	456	16	1.4	3787	22	AAH81775	Human differential
c 384	16	1.4	2274	16	AAQ75959	c 457	16	1.4	3826	22	AAH54225	S. epidermidis gen
c 385	16	1.4	2288	22	ABL10501	458	16	1.4	3891	23	ABL24402	Drosophila melanog
c 386	16	1.4	2322	22	ABL21072	c 459	16	1.4	3967	23	ABL07526	Drosophila melanog
c 387	16	1.4	2340	20	AA562676	c 460	16	1.4	3982	23	ABL23044	Drosophila melanog
c 388	16	1.4	2354	21	AA550348	c 461	16	1.4	4010	23	ABL18552	Drosophila melanog
c 389	16	1.4	2393	21	AAZ46135	c 462	16	1.4	4265	22	AAK52361	Human polynucleoti
c 390	16	1.4	2433	22	AD09336	463	16	1.4	4295	16	AA158456	Human polynucleoti
c 391	16	1.4	2450	23	ABL25910	464	16	1.4	4351	17	AA156642	Coturnix sp. aroma
c 392	16	1.4	2456	22	AAO13362	465	16	1.4	4379	22	AAK53395	Human polynucleoti
c 393	16	1.4	2540	22	AAH14650	466	16	1.4	4379	22	AA160242	Human polynucleoti
c 394	16	1.4	2565	8	AAH70548	c 467	16	1.4	4409	21	AAV6585	Nucleotide sequenc
c 395	16	1.4	2565	8	AAH70548	c 468	16	1.4	4409	21	AAV6585	Nucleotide sequenc
c 396	16	1.4	2565	8	AAH70548	c 469	16	1.4	4409	22	AAV6585	Nucleotide sequenc
c 397	16	1.4	2565	8	AAH70548	c 470	16	1.4	4409	22	AAV6585	Nucleotide sequenc
c 398	16	1.4	2619	23	AA571968	c 471	16	1.4	4409	22	AAV6585	Nucleotide sequenc
c 399	16	1.4	2632	23	ABU04241	c 472	16	1.4	4425	18	AA172781	Human interleukin-
c 400	16	1.4	2639	21	AA593040	c 473	16	1.4	4448	23	ABL14412	Drosophila melanog
c 401	16	1.4	2646	15	AAQ61607	c 474	16	1.4	4480	20	AAAX13257	Enterococcus faeca
c 402	16	1.4	2647	12	AAQ09339	c 475	16	1.4	4534	20	AAAX3042	Human IL-1ra BAC c
c 403	16	1.4	2678	23	ABU10806	476	16	1.4	4539	22	ABA07240	Human pancreatic c
c 404	16	1.4	2679	23	AA594559	477	16	1.4	4539	22	AAK89881	Human digestive sy
c 405	16	1.4	2701	24	ABU34325	c 478	16	1.4	4600	23	ABL18056	Drosophila melanog
c 406	16	1.4	2733	22	AAH14335	c 479	16	1.4	4647	23	ABL07574	Drosophila melanog
c 407	16	1.4	2763	21	AAH65286	c 480	16	1.4	4692	23	ABL13402	Drosophila melanog
c 408	16	1.4	2805	22	AA563223	c 481	16	1.4	4702	23	ABL10820	Drosophila melanog
c 409	16	1.4	2832	23	ABU04243	482	16	1.4	4708	22	AAH18684	Human CDNA sequenc
c 410	16	1.4	2915	22	AA511662	483	16	1.4	4767	21	AAV71571	S. aggregatum PKS
c 411	16	1.4	2947	16	AAH87269	c 484	16	1.4	4968	20	AAK20597	Polynucleotide seq
c 412	16	1.4	2950	22	AAH17486	c 485	16	1.4	4980	22	AAH16236	Mouse ATP-binding
c 413	16	1.4	3022	22	ABU09338	486	16	1.4	5017	23	AA592633	DNA encoding novel
c 414	16	1.4	3051	23	ABU23607	487	16	1.4	5094	23	AA578835	DNA encoding novel
c 415	16	1.4	3095	16	AAQ93911	488	16	1.4	5433	22	AAH99527	Human protein enco
c 416	16	1.4	3099	23	AA553411	c 489	16	1.4	5541	23	ABL18092	Drosophila melanog
c 417	16	1.4	3109	16	AAQ93901	c 490	16	1.4	5767	23	ABL14646	Drosophila melanog
c 418	16	1.4	3118	21	AAH15989	c 491	16	1.4	6048	19	AAV09029	Human hml sodium c
c 419	16	1.4	3130	22	AA571971	c 492	16	1.4	6048	22	AA571971	Human SCNA gene.
c 420	16	1.4	3133	22	AAQ09340	c 493	16	1.4	6063	22	AA546338	Tumour suppressor
c 421	16	1.4	3145	22	AAH75043	c 494	16	1.4	6074	23	ABL06630	Drosophila melanog
c 422	16	1.4	3145	22	AAH75043	c 495	16	1.4	6117	23	ABL34491	Human metastasis a
c 423	16	1.4	3157	23	AAH54685	c 496	16	1.4	6120	23	ABL01860	Drosophila melanog
c 424	16	1.4	3175	23	AA594564	c 497	16	1.4	6200	22	AA546442	Tumour suppressor
c 425	16	1.4	3182	21	AA586611	c 498	16	1.4	6259	24	ABL32290	Human immune syste
c 426	16	1.4	3195	23	AA587692	c 499	16	1.4	6326	24	ABL22338	Chemically treated
c 427	16	1.4	3209	14	AAQ34779	c 500	16	1.4	6664	24	AA561368	Human gene regulat
c 428	16	1.4	3269	23	ABU21122	c 501	16	1.4	6911	17	AA534177	Corynebacterium
c 429	16	1.4	3286	22	AAH54271	c 502	16	1.4	7030	23	ABL12034	Drosophila melanog
c 430	16	1.4	3296	21	AA576408	c 503	16	1.4	7238	24	ABL32448	Human immune syste
c 431	16	1.4	3301	21	AA551259	c 504	16	1.4	7259	23	ABL27520	Drosophila melanog
c 432	16	1.4	3301	21	AA551259	c 505	16	1.4	7306	24	ABL33636	Human immune syste
c 433	16	1.4	3316	21	AA572822	c 506	16	1.4	7488	24	ABL33113	Human immune syste
c 434	16	1.4	3316	21	AA572822	c 507	16	1.4	7526	22	AA540044	Genomic sequence #
c 435	16	1.4	3316	22	AA521482	c 508	16	1.4	7626	22	AA521482	Human digestive sy
c 436	16	1.4	3317	22	AA521482	c 509	16	1.4	7831	22	ABA19111	Human nervous syst
c 437	16	1.4	3351	22	AA521482	c 510	16	1.4	7937	22	AA521482	Human musculoskele
c 438	16	1.4	3355	20	AA521482	c 511	16	1.4	7980	19	AA521482	Lettuce resistance
c 439	16	1.4	3355	20	AA521482	c 512	16	1.4	8530	22	AA521482	Human diagnostic a
c 440	16	1.4	3355	21	AA521482	c 513	16	1.4	8639	20	AA521482	Human IL-1ra BAC c
c 441	16	1.4	3355	21	AA521482	c 514	16	1.4	8777	23	ABL23606	Drosophila melanog
c 442	16	1.4	3364	22	AA521482	c 515	16	1.4	8876	24	ABL34076	Human immune syste
c 443	16	1.4	3384	22	AA521482	c 516	16	1.4	9021	22	AA521482	Human immune syste
c 444	16	1.4	3413	22	AA521482	c 517	16	1.4	9129	22	AA521482	Human immune syste
c 445	16	1.4	3427	22	AA521482	c 518	16	1.4	9151	22	ABA14677	Human genomic DNA
c 446	16	1.4	3442	22	AA521482	c 519	16	1.4	9153	22	ABA14677	Human nervous syst
c 447	16	1.4	3449	21	AA521482	c 520	16	1.4	9254	22	AA521482	Chemically pretrea

c 521	16	1.4	9254	22	AAS46392	Tumour suppressor
522	16	1.4	9295	23	ABL06144	Drosophila melanog
523	16	1.4	9295	23	ABL07236	Drosophila melanog
524	16	1.4	9664	24	ABL32099	Human immune syste
c 525	16	1.4	10078	22	AAS40045	Genomic sequence #
c 526	16	1.4	10078	22	AAK91462	Human digestive sy
c 527	16	1.4	10088	22	AAK5072	Human immune/haema
c 528	16	1.4	10169	14	AAO53522	Cellulose synthase
c 529	16	1.4	10191	23	ABL04240	Drosophila melanog
c 530	16	1.4	10283	23	ABL04242	Drosophila melanog
c 531	16	1.4	10655	23	ABL15398	Drosophila melanog
c 532	16	1.4	10929	22	AAS32755	Human genomic DNA
c 533	16	1.4	11172	22	AAH41188	Murine oligonucleo
c 534	16	1.4	11206	21	AAZ47807	Vector for trappin
c 535	16	1.4	11376	23	ABL11258	Drosophila melanog
c 536	16	1.4	11710	24	AAH43718	E. coli genome sec
c 537	16	1.4	11711	22	ABA18962	Human nervous syst
c 538	16	1.4	11713	22	ABA18963	Human nervous syst
c 539	16	1.4	11881	22	AAS36624	Human cardiovascular
540	16	1.4	11881	22	AAS27691	DNA encoding novel
541	16	1.4	11881	22	AAS27693	DNA encoding novel
542	16	1.4	12831	22	AAK70373	Human immune/haema
c 543	16	1.4	13036	23	ABL10282	Drosophila melanog
544	16	1.4	13154	20	AAK13275	Enterococcus faeca
c 545	16	1.4	13537	22	AAK68964	Human immune/haema
c 546	16	1.4	13574	22	AAK72319	Human immune/haema
c 547	16	1.4	13865	20	AAK13137	Enterococcus faeca
c 548	16	1.4	13982	22	AAK65456	Human immune/haema
c 549	16	1.4	13982	22	AAK77762	Human immune/haema
c 550	16	1.4	13982	22	AAK84032	Human immune/haema
c 551	16	1.4	14904	22	AAK75386	Human immune/haema
c 552	16	1.4	15090	22	AAK72319	Human immune/haema
c 553	16	1.4	15254	22	AAK83135	Human immune/haema
c 554	16	1.4	15254	22	AAK83137	Human immune/haema
c 555	16	1.4	15255	22	AAK83136	Human immune/haema
c 556	16	1.4	15255	22	AAK83134	Human immune/haema
557	16	1.4	15366	22	AAS28623	Genomic sequence #
c 558	16	1.4	16159	22	ABA18491	Human nervous syst
c 559	16	1.4	16423	22	AAAD16645	Human novel protei
560	16	1.4	16423	22	AAS30242	DNA encoding rena
c 561	16	1.4	16748	22	AAK65193	Human immune/haema
c 562	16	1.4	16842	22	AAS46412	Tumour suppressor
c 563	16	1.4	16842	22	AAK61336	Human gene regulat
c 564	16	1.4	16870	20	AAK13035	Enterococcus faeca
c 565	16	1.4	17934	24	ABL93319	Human immune syste
c 566	16	1.4	18085	23	ABL09930	Drosophila melanog
c 567	16	1.4	18385	22	AAS28584	Genomic sequence #
c 568	16	1.4	18385	22	AAK79804	Human immune/haema
569	16	1.4	18786	23	AAS59579	Protonibacterium
570	16	1.4	19231	21	AAAB1470	N. meningitidis pa
c 571	16	1.4	19619	22	AAK28527	Genomic fragment #
572	16	1.4	20645	22	AAI05355	Human reproductiv
573	16	1.4	20978	23	ABL20786	Drosophila melanog
c 574	16	1.4	21835	22	AAK66483	Human immune/haema
c 575	16	1.4	21913	22	AAI03377	Human reproductiv
c 576	16	1.4	21913	22	AAI03378	Human reproductiv
c 577	16	1.4	21913	22	AAI03379	Human reproductiv
c 578	16	1.4	23075	22	ABA19112	Human nervous syst
c 579	16	1.4	24417	18	AAK79221	Pseudomonas aerugi
c 580	16	1.4	26006	22	AAI99233	Human excretory re
c 581	16	1.4	26006	22	AAS36463	Human cardiovascular
c 582	16	1.4	26006	22	AAI03357	Human reproductiv
c 583	16	1.4	26006	22	AAS26664	Human genomic DNA
c 584	16	1.4	26006	22	AAI62592	Human breast or ov
c 585	16	1.4	26006	22	AAI63583	Human kidney relat
c 586	16	1.4	26013	22	AAI99231	Human excretory re
c 587	16	1.4	26013	22	AAS36461	Human cardiovascular
c 588	16	1.4	26013	22	AAI03355	Human reproductiv
589	16	1.4	26013	22	AAS26662	Human genomic DNA
c 590	16	1.4	26013	22	AAI62590	Human breast or ov
c 591	16	1.4	26018	22	AAI63581	Human kidney relat
c 592	16	1.4	26018	22	AAI99232	Human excretory re
c 593	16	1.4	26018	22	AAS36462	Human cardiovascular
c 594	16	1.4	26018	22	AAI03596	Human reproductiv
c 595	16	1.4	26018	22	AAS26663	Human genomic DNA
c 596	16	1.4	26018	22	AAI62591	Human breast or ov
c 597	16	1.4	26018	22	AAI63582	Human kidney relat
c 598	16	1.4	27579	22	AAK66484	Human immune/haema
c 599	16	1.4	32170	22	AAS28674	Genomic sequence #
c 600	16	1.4	32249	22	ABL04676	Human reproductiv
c 601	16	1.4	36445	23	ABL18116	Drosophila melanog
c 602	16	1.4	36501	22	AAK64829	Human immune/haema
c 603	16	1.4	37337	23	AAK59518	Protonibacterium
604	16	1.4	49999	20	AAZ23901	Human LOBO homolog
605	16	1.4	50000	21	AAK6364	Polymorphic repeat
606	16	1.4	68750	21	AAZ55887	Sorangium cellulos
c 607	16	1.4	69936	21	AAAB1479	N. meningitidis pa
c 608	16	1.4	71989	21	AAK29349	Sorangium cellulos
609	16	1.4	80251	23	ABL16442	Drosophila melanog
610	16	1.4	80251	23	ABL16448	Drosophila melanog
611	16	1.4	83390	21	AAK22283	BAC containing rep
612	16	1.4	90336	21	AAK22289	BAC containing rep
613	16	1.4	95223	21	AAK22282	BAC containing rep
c 614	16	1.4	99960	21	AAZ50905	Human RBC-1 partia
615	16	1.4	107820	22	AAAD16230	Human ATP-binding
c 616	16	1.4	124884	22	AAH74201	Nucleotide sequenc
c 617	16	1.4	125157	22	AAH74202	Nucleotide sequenc
618	16	1.4	172325	21	AAK21613	Neisseria meningit
619	16	1.4	335913	22	AAI61371	Soybean 240017 reg
620	16	1.4	335913	22	AAI61372	Soybean 240017 reg
c 621	16	1.4	349980	22	AAH64966	C glutaminc codin
622	16	1.4	349980	22	AAH68528	C glutaminc codin
c 623	16	1.4	465237	24	ABA90193	Human oestrogen re
c 624	16	1.4	465237	24	AAK22303	Arabidopsis thalia
c 625	16	1.4	611590	21	AAK95240	Human neuregulin-1
c 626	16	1.4	1503900	22	AAK67733	Human neuregulin-1
c 627	16	1.4	1664976	19	AAK21209	Methanococcus jan
c 628	16	1.4	1830121	17	AAK42063	Haemophilus influe
c 629	16	1.4	2944528	24	ABA903041	Listeria monocytog
630	16	1.4	19	21	AAK85195	Cyclin G1 ribozyme
631	16	1.3	19	21	AAK85196	Cyclin G1 ribozyme
632	16	1.3	19	22	AAH60357	Cyclin G1 ribozyme
633	16	1.3	19	22	AAH60358	Cyclin G1 ribozyme
634	16	1.3	23	21	AAK71038	Single nucleotide
635	16	1.3	23	21	AAK71077	Single nucleotide
636	16	1.3	23	21	AAK71089	Single nucleotide
637	16	1.3	23	21	AAK71110	Single nucleotide
638	16	1.3	29	20	AAV99736	Human secreted pro
c 639	16	1.3	30	22	AAAD12655	Human AHC_2H01 CDN
c 640	16	1.3	30	22	AAK88740	Human catenin-bind
c 641	16	1.3	38	22	AAH96137	Human CHK1 ribozym
c 642	16	1.3	49	18	AAI80460	Hepatoma AS-30D Ty
c 643	16	1.3	51	22	AAI79445	Primer (+8) for An
c 644	16	1.3	60	19	AAV49463	Human genome fragm
645	16	1.3	71	15	AAO77161	Staphylococcus aur
c 646	16	1.3	94	23	AAS50348	Human gene express
c 647	16	1.3	104	20	AAZ14021	Human breast cell
648	16	1.3	118	22	ABA48167	Human foetal liver
649	16	1.3	118	22	ABA66044	Human foetal liver
650	16	1.3	118	22	ABA33127	Probe #11593 for g
651	16	1.3	118	22	AAK14467	Human brain expres
652	16	1.3	118	22	AAK42027	Human bone marrow
653	16	1.3	118	22	AAI20975	Probe #10908 for g
654	16	1.3	118	22	AAI6222	Probe #6679 used t
655	16	1.3	118	22	AAI06688	Human immune/haema
656	16	1.3	140	22	AAK57447	Human breast cell
657	16	1.3	141	22	ABA65705	Human foetal liver
658	16	1.3	141	22	ABA65593	Human foetal liver
659	16	1.3	141	22	ABA65593	Human foetal liver
660	16	1.3	141	22	ABA32792	Probe #11258 for g
661	16	1.3	141	22	ABA33660	Probe #12126 for g
662	16	1.3	141	22	AAK14110	Human brain expres
663	16	1.3	141	22	AAK15020	Human brain expres
664	16	1.3	141	22	AAK39852	Human bone marrow
665	16	1.3	141	22	AAK40751	Human bone marrow
c 666	16	1.3	141	22	AAI20662	Probe #10595 for g

c 667	15	1.3	141	22	AA121514	Probe #11447 for g	c 740	15	1.3	380	14	AA059718	Human brain Express
c 668	15	1.3	141	22	AA145873	Probe #14559 used	c 741	15	1.3	380	22	AA527582	CDNA encoding nove
c 669	15	1.3	141	22	AA146806	Probe #15492 used	c 742	15	1.3	381	23	AA564398	DNA encoding novel
c 670	15	1.3	141	22	AA106362	Probe #6353 used t	c 743	15	1.3	383	21	AA667080	Pinus radiata cell
c 671	15	1.3	141	22	AA107210	Probe #7201 used t	c 744	15	1.3	383	22	ABA11395	Human nervous syst
c 672	15	1.3	148	22	ABA8694	Human breast cell	c 745	15	1.3	385	22	AA136640	Probe #5326 used t
c 673	15	1.3	148	22	AA121530	Probe #11463 for g	c 746	15	1.3	387	19	AAV44958	Anti-HIV-1 group O
c 674	15	1.3	148	22	AA107226	Probe #7217 used t	c 747	15	1.3	388	22	AA541553	CDNA encoding nove
c 675	15	1.3	159	17	AA131610	Klebsiella pneumonia	c 748	15	1.3	389	21	AA598827	Human proliferatio
c 676	15	1.3	160	23	AA549197	Staphylococcus aur	c 749	15	1.3	389	21	AA598108	Human breast cance
c 677	15	1.3	165	18	AAV75450	Staphylococcus aur	c 750	15	1.3	389	22	AA125952	Human breast cell
c 678	15	1.3	160	22	AAH64969	C glutamylam codin	c 751	15	1.3	390	22	ABA45095	Human breast cell
c 679	15	1.3	191	21	AAAC12991	Human secreted pro	c 752	15	1.3	390	22	AAK03802	Human brain expres
c 680	15	1.3	195	22	AA126235	Human breast cance	c 753	15	1.3	390	22	AAK29269	Human bone marrow
c 681	15	1.3	199	22	AA119185	Human breast cance	c 754	15	1.3	390	22	AA13868	Probe #3801 for ge
c 682	15	1.3	216	17	AA131609	Klebsiella pneumonia	c 755	15	1.3	390	22	AA135229	Probe #3915 used t
c 683	15	1.3	217	21	AACT16279	Human secreted pro	c 756	15	1.3	390	22	AA103772	Probe #3723 used t
c 684	15	1.3	226	11	AAO05051	Recombinant AatI g	c 757	15	1.3	392	22	AAH55612	Human breast tumor
c 685	15	1.3	250	23	AA549843	Staphylococcus aur	c 758	15	1.3	395	21	AAH87152	Rat hepatocyte car
c 686	15	1.3	250	23	AA549849	Staphylococcus aur	c 759	15	1.3	395	22	AA183339	Human polynucleoti
c 687	15	1.3	257	18	AAAT72461	Human CII-3 gene 1	c 760	15	1.3	396	22	AB124809	Drosophila melanog
c 688	15	1.3	252	16	AAAT23232	Human gene signatu	c 761	15	1.3	397	22	AA655942	Novel human polyn
c 689	15	1.3	262	21	AAAC24397	Human secreted pro	c 762	15	1.3	399	22	ABA54022	Human foetal liver
c 690	15	1.3	263	19	AAV03373	2-Oxo-glutarate de	c 763	15	1.3	399	22	ABA23776	Probe #2242 for ge
c 691	15	1.3	268	21	AAAT70684	Human alpha-Gal A	c 764	15	1.3	399	22	AAK02295	Human brain expres
c 692	15	1.3	268	21	AAA75433	5' end of cDNA of	c 765	15	1.3	399	22	AAK27739	Human bone marrow
c 693	15	1.3	287	13	AAO26399	PCIB4223. Synthet	c 766	15	1.3	399	22	AA112319	Probe #2252 for ge
c 694	15	1.3	289	20	AAV88814	EST clone HO1216.	c 767	15	1.3	399	22	AA133675	Probe #2261 used t
c 695	15	1.3	284	20	AAV41489	Human secreted pro	c 768	15	1.3	399	22	AA102233	Probe #2224 used t
c 696	15	1.3	287	21	AAAC04510	Human secreted pro	c 769	15	1.3	401	21	AAZ80677	Human colon cancer
c 697	15	1.3	298	22	AA135287	Human musculoskele	c 770	15	1.3	401	22	AAK96571	Human neurogulin g
c 698	15	1.3	300	20	AA214557	Human gene express	c 771	15	1.3	401	22	AAK96573	Human neurogulin g
c 699	15	1.3	300	20	AA212676	Human gene express	c 772	15	1.3	401	22	AAK98064	Human neurogulin g
c 700	15	1.3	303	20	AAV89591	EST clone CN951.	c 773	15	1.3	401	22	AAK98066	Human neurogulin g
c 701	15	1.3	308	22	AA136581	Human musculoskele	c 774	15	1.3	402	14	AA037826	Sequence of the pa
c 702	15	1.3	309	22	AA190840	Human polynucleoti	c 775	15	1.3	403	22	AA673782	Novel human polyn
c 703	15	1.3	310	20	AAV40636	Human secreted pro	c 776	15	1.3	403	15	AAQ44373	Sequence of VHH DN
c 704	15	1.3	310	22	AA124510	Human breast cance	c 777	15	1.3	405	22	ABA08492	Human taste recept
c 705	15	1.3	312	21	AAAC04313	Human secreted pro	c 778	15	1.3	410	23	ABU03565	Drosophila melanog
c 706	15	1.3	315	22	AA135120	Human musculoskele	c 779	15	1.3	415	20	AAV06971	Zs19g expressed se
c 707	15	1.3	315	22	AA100773	Human reproductive	c 780	15	1.3	415	21	AAZ08285	Expressed Sequence
c 708	15	1.3	320	20	AAV98198	Nucleotide sequenc	c 781	15	1.3	415	22	AA527352	CDNA encoding nove
c 709	15	1.3	322	22	AA539155	Novel human diagno	c 782	15	1.3	416	21	AAAC24897	Human secreted pro
c 710	15	1.3	327	21	AA551986	Salmonella Pathoge	c 783	15	1.3	418	22	AA527817	DNA encoding novel
c 711	15	1.3	327	22	AAK87158	Human immune/haema	c 784	15	1.3	421	21	AAAC92324	ADP-ribosylation f
c 712	15	1.3	330	20	AAK41237	Human secreted pro	c 785	15	1.3	421	22	AAAC32214	Human secreted pro
c 713	15	1.3	332	22	ABA75299	Human foetal liver	c 786	15	1.3	421	22	ABA57132	Human foetal liver
c 714	15	1.3	332	22	ABA39952	Probe #18418 for g	c 787	15	1.3	425	22	AAH11765	Human CDNA clone (
c 715	15	1.3	332	22	AA539159	Novel human diagno	c 788	15	1.3	426	22	AAK54082	Murine transcripti
c 716	15	1.3	332	22	AAK23852	Human brain expres	c 789	15	1.3	427	22	AA67464	Novel human polyn
c 717	15	1.3	332	22	AAK49933	Human bone marrow	c 790	15	1.3	430	22	ABA88798	Escherichia coli p
c 718	15	1.3	332	22	AA127030	Probe #16963 for g	c 791	15	1.3	431	22	AA108207	Human breast cance
c 719	15	1.3	332	22	AA155863	Probe #24549 used	c 792	15	1.3	433	22	AA112678	Human breast cance
c 720	15	1.3	334	21	AAAC29477	Human secreted pro	c 793	15	1.3	435	22	AAH52461	S. epidermidis ope
c 721	15	1.3	339	22	AA589368	DNA encoding novel	c 794	15	1.3	441	20	AAK83998	Conopeptide propep
c 722	15	1.3	347	21	AACT1070	Single nucleotide	c 795	15	1.3	441	22	AAK70369	Human immune/haema
c 723	15	1.3	347	21	AACT1109	Single nucleotide	c 796	15	1.3	444	22	AA198964	Human excretory re
c 724	15	1.3	348	22	AAK64288	Human immune/haema	c 797	15	1.3	444	22	AA163314	Human kidney relat
c 725	15	1.3	348	22	AA535958	Human cardiovascular	c 798	15	1.3	449	22	AA163314	Human kidney relat
c 726	15	1.3	348	22	AA535959	Human cardiovascular	c 799	15	1.3	450	19	AAV48495	Amplicon DNA compr
c 727	15	1.3	348	22	AAK03578	Cynomolgus IAP-11k	c 800	15	1.3	450	22	ABA13905	Anti-HIV-1 group O
c 728	15	1.3	351	22	AAK03576	Baboon IAP-1 like pr	c 801	15	1.3	452	22	ABA18305	Human nervous syst
c 729	15	1.3	351	22	AAK03579	Gorilla IAP-1 like p	c 802	15	1.3	455	22	ABA42688	Human nervous syst
c 730	15	1.3	351	22	AAK03580	Rhesus IAP-1 like pr	c 803	15	1.3	455	22	ABA53112	Human breast cell
c 731	15	1.3	352	22	AA186079	Human polynucleoti	c 804	15	1.3	455	22	ABA22886	Human foetal liver
c 732	15	1.3	362	21	AAAC26554	Human secreted pro	c 805	15	1.3	455	22	AAK01362	Probe #1352 for ge
c 733	15	1.3	365	21	AACT1040	Single nucleotide	c 806	15	1.3	455	22	AAK26818	Human brain expres
c 734	15	1.3	365	21	AACT1079	Single nucleotide	c 807	15	1.3	455	22	AA111448	Human bone marrow
c 735	15	1.3	365	21	AACT1091	Single nucleotide	c 808	15	1.3	455	22	AA132719	Probe #1381 for ge
c 736	15	1.3	365	21	AACT1112	Single nucleotide	c 809	15	1.3	455	22	AA101365	Probe #1405 used t
c 737	15	1.3	366	22	ABA01254	Human Erbin probe	c 810	15	1.3	456	21	AA59434	Probe #1356 used t
c 738	15	1.3	367	23	AA592301	DNA encoding novel	c 811	15	1.3	461	22	AA119117	Nucleotide sequenc
c 739	15	1.3	369	22	AAH93990	Human foetal CDNA,	c 812	15	1.3	465	23	AA564971	Human encoding novel

813	1.3	466	22	ABA56596	Human foetal liver	c 886	15	1.3	595	22	ABA06665	Human cDNA SEQ ID
814	1.3	471	22	ABA43040	Human breast cell	c 887	15	1.3	595	22	AAS28835	Human immunoglobulin
815	1.3	471	22	ABA53454	Human foetal liver	c 888	15	1.3	596	22	ABA62787	Human foetal liver
816	1.3	471	22	ABA23225	Probe #1691 for ge	889	15	1.3	596	22	ABA30080	Probe #8546 for ge
817	1.3	471	22	AAK01724	Human brain expres	890	15	1.3	596	22	AAK11185	Human brain expres
818	1.3	471	22	AAK27177	Human bone marrow	891	15	1.3	596	22	AAK35992	Human bone marrow
819	1.3	471	22	AAI11763	Probe #1696 for ge	892	15	1.3	596	22	AAI17834	Probe #11490 used
820	1.3	471	22	AAI33073	Probe #1759 used t	893	15	1.3	596	22	AAI42804	Probe #17767 for ge
821	1.3	471	22	AAI01692	Probe #1683 used t	894	15	1.3	598	22	ABA63325	Human foetal liver
822	1.3	472	22	ABA59318	Human foetal liver	895	15	1.3	598	22	AAI18329	Probe #8262 for ge
823	1.3	472	22	AAK07536	Human brain expres	c 896	15	1.3	607	17	AAT26893	Haemophilus influe
824	1.3	472	22	AAK33332	Human bone marrow	c 897	15	1.3	607	22	AAK83921	Human immune/haema
825	1.3	472	22	AAI39119	Probe #7805 used t	c 898	15	1.3	608	22	AAS41002	cDNA encoding nove
826	1.3	473	22	AAK57817	Human immune/haema	c 899	15	1.3	612	22	AAH31452	Human olfactory re
827	1.3	474	18	AAT67779	H. pylori membrane	900	15	1.3	614	22	AAH31450	Human olfactory re
828	1.3	474	18	AAT77459	H. pylori transmem	c 901	15	1.3	618	22	AAK78164	Human immune/haema
829	1.3	474	22	AAS09369	Amplicon DNA compr	c 902	15	1.3	618	22	AAK78165	Human immune/haema
830	1.3	475	21	AAAS6026	Human chromosome 1	c 903	15	1.3	618	22	AAK78166	Human immune/haema
831	1.3	475	22	AAK65082	Human immune/haema	904	15	1.3	622	19	AAT98746	DNA encoding a S.
832	1.3	480	21	AAAC00962	Human secreted pro	905	15	1.3	628	22	AAH31454	Human olfactory re
833	1.3	485	21	AAV86871	EST clone BK200.	c 906	15	1.3	629	21	AAH31453	Aspergillus oryzae
834	1.3	485	21	AAAB2372	N. meningitidis pa	c 907	15	1.3	632	22	AAAF12979	Human olfactory re
835	1.3	489	20	AAX98196	Nucleotide sequenc	c 908	15	1.3	634	22	AAK71202	Human immune/haema
836	1.3	489	22	ABA43584	Human breast cell	909	15	1.3	635	18	AAV74505	Human immune/haema
837	1.3	489	22	AAI12335	Probe #2268 for ge	910	15	1.3	643	22	AAH07292	Human cDNA clone (
838	1.3	489	22	AAI02249	Probe #2240 used t	c 911	15	1.3	647	20	AAH37381	Human secreted pro
839	1.3	492	21	AAZ53095	Neisseria gonorrhe	912	15	1.3	649	20	AAH37381	Human secreted pro
840	1.3	496	22	AAK58660	Human immune/haema	913	15	1.3	649	20	AAH37381	Human secreted pro
841	1.3	500	21	AAAC3588	Cat flea head and	914	15	1.3	649	20	AAH37381	Human secreted pro
842	1.3	501	23	AAAS53711	Helicobacter pylori	915	15	1.3	651	19	AAV19119	Human secretory pe
843	1.3	501	24	AAAS61795	Lung small cell ca	916	15	1.3	651	20	AAV21557	DNA encoding a Str
844	1.3	509	19	AAI14418	H. pylori GHPD 693	917	15	1.3	653	22	AAAS6387	S. pneumoniae M pr
845	1.3	510	16	AAO92284	Human CTLA-8 cDNA.	c 918	15	1.3	653	23	AAAS49439	Human cDNA for an
846	1.3	510	24	AAH77244	Human cytokine-lik	c 919	15	1.3	653	23	AAAS4950	Staphylococcus aur
847	1.3	514	21	AAAC78252	Human cancer assoc	c 920	15	1.3	654	22	AAAO9113	Staphylococcus aur
848	1.3	514	22	AAH64923	Human secreted pro	921	15	1.3	657	21	AAFO7813	Human breast cance
849	1.3	517	22	ABA63078	Human foetal liver	922	15	1.3	663	22	ABA89116	Rusarium venenatum
850	1.3	517	22	ABA30335	Probe #8799 for ge	923	15	1.3	667	22	ABA18844	Escherichia coli p
851	1.3	517	22	AAK11502	Human brain expres	924	15	1.3	667	22	ABA18844	Human nervous syst
852	1.3	517	22	AAK37280	Human bone marrow	c 925	15	1.3	667	22	AAH29009	Drosophila melanog
853	1.3	517	22	AAI18109	Probe #8042 for ge	926	15	1.3	669	22	AAH68159	C glutamicum codin
854	1.3	517	22	AAI43113	Probe #11799 used	c 927	15	1.3	677	18	AAV75974	Staphylococcus aur
855	1.3	519	22	ABA60843	Human foetal liver	928	15	1.3	677	21	AAAC39805	Arabidopsis thalia
856	1.3	519	22	ABA28850	Probe #7316 for ge	929	15	1.3	678	22	AAK92248	Human cDNA 5'-end
857	1.3	519	22	AAK09130	Human brain expres	930	15	1.3	678	22	AAK93508	Human calcium chan
858	1.3	519	22	AAK35019	Human bone marrow	931	15	1.3	681	13	AAO29259	Human cDNA clone r
859	1.3	519	22	AAI40736	Probe #9422 used t	932	15	1.3	684	20	AAI20377	Borrelia burgdorfe
860	1.3	536	21	AAA68132	Pinus radiata flav	933	15	1.3	689	22	AAI24694	Human breast cance
861	1.3	537	17	AAI31608	Klebsiella pneumon	934	15	1.3	695	22	AAAS9843	DNA encoding a tag
862	1.3	546	21	AAZ38326	Human transmembran	c 935	15	1.3	706	22	AAK87618	Human immune/haema
863	1.3	555	21	AAAC41592	Arabidopsis thalia	c 936	15	1.3	706	22	AAK87618	Human immune/haema
864	1.3	555	22	AAH09178	Human cDNA clone (	937	15	1.3	709	19	AAV27371	Escherichia coli p
865	1.3	561	23	ABLI27217	Drosophila melanog	938	15	1.3	711	18	AAT67962	H. pylori transmem
866	1.3	563	22	AAK73488	Human immune/haema	c 939	15	1.3	712	21	AAAC36852	Arabidopsis thalia
867	1.3	563	22	AAK80627	Human immune/haema	940	15	1.3	714	20	AAK97884	Human secreted pro
868	1.3	564	24	ABK09691	Human ovarian tumo	c 941	15	1.3	717	22	AAK87616	Human immune/haema
869	1.3	565	18	AAFI0630	Rusarium venenatum	942	15	1.3	718	20	AAZ10655	cDNA encoding a hu
870	1.3	567	21	AAV13869	Gene coding for L-	943	15	1.3	719	23	ABLI19255	Drosophila melanog
871	1.3	567	21	AAFO8472	Fusarium venenatum	c 944	15	1.3	720	14	AAO45918	Phage PI ref gene.
872	1.3	570	22	AAH08763	Human cDNA clone (	945	15	1.3	722	22	AAH04090	Human cDNA clone (
873	1.3	572	19	AAI14458	H. pylori GHPD 951	946	15	1.3	722	22	AAK75721	Human immune/haema
874	1.3	574	22	ABA60564	Human foetal liver	947	15	1.3	723	22	AAK61799	Human immune/haema
875	1.3	574	22	ABA28708	Probe #7174 for ge	c 948	15	1.3	726	23	AAAS73558	DNA encoding novel
876	1.3	574	22	AAK08846	Human brain expres	949	15	1.3	729	22	AAH08057	Human cDNA clone (
877	1.3	574	22	AAK34734	Human bone marrow	950	15	1.3	730	17	AAAT88425	K. pneumoniae dete
878	1.3	574	22	AAI16821	Probe #6754 for ge	951	15	1.3	730	22	ABA76835	Human polynucleoti
879	1.3	577	22	AAI10453	Probe #9139 used t	952	15	1.3	733	22	AAK53226	Aspergillus oryzae
880	1.3	577	22	AAI15785	Human breast cance	953	15	1.3	733	21	AAFI12207	Human musculoskele
881	1.3	579	22	AAK11759	Human brain expres	c 954	15	1.3	735	22	AAI136929	S. epidermidis ope
882	1.3	579	23	AAH09463	Human cDNA clone (	955	15	1.3	743	22	AAK59603	Human immune/haema
883	1.3	579	23	AAAS70599	DNA encoding novel	956	15	1.3	743	22	AAK59603	Human breast cance
884	1.3	580	23	AAAS87180	DNA encoding novel	957	15	1.3	752	22	AAI24658	Human cDNA clone (
885	1.3	594	22	AAAS02054	DNA encoding molec	958	15	1.3	756	22	AAH03408	Human cDNA clone (



QY 721 tatttaacaaacggagcgttaagaagaagactgttacgcgcagccagatgcatctgcg 780  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 tatttaacaaacggagcgttaagaagaagactgttacgcgcagccagatgcatctgcg 780  
 QY 781 ggtggtgcatctactcgtatgaacaacccgtgaagatgatgtatgattgtgtgcag 840  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 ggtggtgcatctactcgtatgaacaacccgtgaagatgatgtatgattgtgtgcag 840  
 QY 841 acgttaacatacagcagactctacgcattctacccgtatgaagaattatagactctac 900  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 841 acgttaacatacagcagactctacgcattctacccgtatgaagaattatagactctac 900  
 QY 901 gcatctatggttaaaagatgattacggtatcatatggtgctacccctgattcccg 960  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 901 gcatctatggttaaaagatgattacggtatcatatggtgctacccctgattcccg 960  
 QY 961 tcacagcgttcaacccgacgataaactgctgttaaatctatcgcgatttcgttaaa 1020  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 961 tcacagcgttcaacccgacgataaactgctgttaaatctatcgcgatttcgttaaa 1020  
 QY 1021 actgaaacaaactgataaactgacgagatggttttcgcgtgtgtctctgtccagcgac 1080  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1021 actgaaacaaactgataaactgacgagatggttttcgcgtgtgtctctgtccagcgac 1080  
 QY 1081 aacgcgcgccataatgaacagctggtcgatcgagctactcgttgaagaacctgcactaa 1137  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1081 aacgcgcgccataatgaacagctggtcgatcgagctactcgttgaagaacctgcactaa 1137  
  
 RESULT 2  
 ID AAF60839 standard; DNA; 2122 BP.  
 AC AAF60839;  
 XX  
 DE 16-MAY-2001 (first entry)  
 XX  
 DE Pseudomonas sp export system associated DNA ORF02378a.  
 XX  
 KW Export system; transgenic plant; translocation; plant growth;  
 KW plant development; plant yield; soil quality; phytoprotection; ds.  
 XX  
 OS Pseudomonas sp.  
 XX  
 PN DE19935105-A1.  
 XX  
 PD 01-FEB-2001.  
 XX  
 PF 27-JUL-1999; 99DE-1035105.  
 XX  
 PR 27-JUL-1999; 99DE-1035105.  
 XX  
 PA (TIGR-) TIGR INST GENOMIC RES.  
 PA (OUTA-) OUTAGEN GMBH.  
 PA (GBF-) GBS BIOTECHNOLOGISCHE FORSCHUNG MBH.  
 PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.  
 PA (MEDI-) MEDIZINISCHE HOCHSCHULE HANNOVER.  
 XX  
 DR WPI: 2001-192471/20.  
 XX  
 PT New DNA encoding a bacterial export system, useful for promoting growth  
 PT and yield of plants -  
 XX  
 PS Claim 3a; Page 11-12; 26pp; German.  
 XX  
 CC This invention describes novel DNA sequences (1) that express products  
 CC having the biological function of export systems. The invention also  
 CC describes (a) recombinant expression vectors containing (1); (b)  
 CC prokaryotic and eukaryotic cells transformed or transfected with (1) or  
 CC the vector of (a); (c) production of export systems by culturing cells of  
 CC (b); (d) (partial) expression products (II) of (1), and synthetic  
 CC proteins or peptides with the same sequences; (e) mono- or poly-clonal  
 CC antibodies (Ab) specific for (II); (f) hybridoma cells that produce

CC monoclonal Ab; and (g) transgenic plants that contain cells of (b). The  
 CC export systems encoded by (1) translocate, through the bacterial cell  
 CC wall, endogenous or modified metabolites or endogenously or  
 CC heterologously expressed proteins. (1), and their fragments, are useful:  
 CC (1) for expression of recombinant export systems; (2) as probes and  
 CC primers for detection, isolation and amplification of full-length cDNA  
 CC sequences; and (3) for producing transgenic plants. Cells transformed  
 CC with (1), also recombinant export systems or synthetic peptides or  
 CC proteins with the same activities, are used to promote growth,  
 CC development and yield of plants, particularly leguminosae. Host cells  
 CC that express export systems improve soil quality, are phytoprotective and  
 CC improve development, growth and yield of plants.  
 CC  
 SQ Sequence 2122 BP; 411 A; 696 C; 705 G; 309 T; 1 other;  
  
 Query Match 1.9%; Score 22; DB 22; Length 2122;  
 Best Local Similarity 100.0%; Pred. No. 0.61;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 399 gctgctgagccttggtggtgag 420  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 2068 gctgctgagccttggtggtgag 2089  
  
 RESULT 3  
 ID ABL01346 standard; DNA; 746 BP.  
 AC ABL01346;  
 XX  
 DE 15-MAR-2002 (first entry)  
 XX  
 DE Murine apoptosis related DNA sequence #11.  
 XX  
 KW Apoptosis; mouse; cancer; autoimmune disease; viral infection;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW reperfusion injury; stroke; liver damage; dilatory cardiomyopathy;  
 KW transgenic animal; hepatotropic; antialcoholism; cytostatic;  
 KW immunosuppressive; virucide; nootropic; neuroprotective; vasotropic;  
 KW antiparkinsonian; cerebroprotective; ds.  
 XX  
 OS Mus sp.  
 XX  
 PN DE10126344-A1.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 30-MAY-2001; 2001DE-1026344.  
 XX  
 PR 14-JUL-2000; 2000DE-1034303.  
 XX  
 PA (PLAC ) MAX PLANCK GRS FORDERUNG WISSENSCHAFTEN.  
 PA Grimm S, Schoenfeld N, Brazillis E, Cramer U, Gewies A, Voss F;  
 PI Mund T, Albayrak T, Galle H, Klein M;  
 XX  
 DR WPI: 2002-115563/16.  
 XX  
 PT New apoptosis-associated nucleic acid sequences and polypeptides,  
 PT useful for diagnosis, treatment and prevention of e.g. tumors and  
 PT neurodegeneration -  
 XX  
 PS Claim 1; Page 136; 227pp; German.  
 XX  
 CC The present invention relates to nucleic acids from the mouse, where the  
 CC nucleic acid is associated with apoptosis. The sequences can be used in  
 CC the diagnosis, treatment and prevention of diseases associated with  
 CC excessive or inadequate apoptosis, including tumours, autoimmune  
 CC diseases, viral infections, degenerative diseases (Alzheimer's,  
 CC Parkinson's and Huntington's diseases), reperfusion injury, stroke and  
 CC alcohol-induced injury to the liver, for identifying agents for treating  
 CC these diseases, and to prepare transgenic animals in which expression of

CC an apoptosis related sequence is altered. These are useful for genetic  
CC and/or pharmacological investigations of apoptosis and related diseases,  
CC including dilatory cardiomyopathy. The present sequence is one of the  
CC apoptosis related sequences of the invention.

XX  
SQ Sequence 746 BP, 123 A; 220 C; 182 G; 201 T; 20 other;

Query Match 1.8%; Score 20; DB 24; Length 746;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 938 tgggcttcaccctgatgtc 957  
|||||  
Db 497 tgggcttcaccctgatgtc 516

## RESULT 4

AAK61604 standard; cDNA; 882 BP.

XX AAK61604;

XX 06-NOV-2001 (first entry)

DE Human Immune/haematopoietic antigen encoding CDNA SEQ ID NO:6664.

XX Human; Immune; haematopoietic; Immune/haematopoietic antigen; Cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ss.

XX Homo sapiens.

PN WO200157182-A2.

XX 09-AUG-2001.

PD 17-JAN-2001; 2001WO-US01354.

XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.





PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139763.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.7%; Score 19; DB 21; length 1282;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 875 ttatgaaagaattacga 893  
|||||  
Db 670 ttatgaaagaattacga 688

RESULT 6  
AAC40124  
ID AAC40124 standard; DNA; 1289 BP.  
XX  
AC AAC40124;  
XX  
DT 17-OCT-2000 (first entry)  
XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27143.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
XX Metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134236.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.

```
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.
Query Match 1.7%; Score 19; DB 21; Length 1289;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 875 ttatgaaagaattataga 893
    |||||||
DB 670 ttatgaaagaattataga 688
```

```
RESULT 7
AAV74939/c
ID AAV74939 standard; DNA: 453 BP.
XX
AC AAV74939;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #628.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
```

```
OS Staphylococcus aureus.
XX
XX EP786519-A2.
XX
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97EP-0100117.
XX
XX 05-JAN-1996; 96US-0009861.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX PI Rosen CA;
XX WPI; 1997-374922/35.
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX PT stored on computer readable medium and used in the production of
XX PT anti-S.aureus vaccines
XX
XX Claim 1; Page 1555; 3271pp; English.
XX
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX homologues of any of the S.aureus DNA sequences contained on the
XX computer readable medium.
XX
XX Sequence 453 BP; 108 A; 82 C; 58 G; 202 T; 3 other;
```

```
OY 566 actctcttgtaacaatgacg 583
    |||||||
DB 178 ACTCTCTTGTGTAATACG 161
Query Match 1.6%; Score 18; DB 18; Length 453;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 8
ABA63976/c
ID ABA63976 standard; DNA: 581 BP.
XX
XX ABA63976;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #12281.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
```

XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX  
XX Claim 1; SEQ ID NO 12281; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;  
SQ

Query Match 1.6%; Score 18; DB 22; Length 581;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ctgtgctgctgctgtct 60  
|||||  
DB 168 CTGTCTGCTGCTGTCT 151

RESULT 9  
ABA31149/c  
ID ABA31149 standard; DNA; 581 BP.  
XX  
AC ABA31149;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Probe #9615 for gene expression analysis in human heart cell sample.  
XX  
KW Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.  
XX  
XX Homo sapiens.  
OS  
PN MO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00666.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488899/53.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
XX Claim 1; SEQ ID No 9615; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart. The  
CC present sequence is one such probe. The probes may be used for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from the human heart via microarrays. By measuring gene expression, the  
CC probes are useful for predicting, diagnosing, grading, staging,  
CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;  
SQ

Query Match 1.6%; Score 18; DB 22; Length 581;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ctgtgctgctgctgtct 60  
|||||  
DB 168 CTGTCTGCTGCTGTCT 151

RESULT 10  
AAK12475/c  
ID AAK12475 standard; DNA; 581 BP.  
XX  
AC AAK12475;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe SEQ ID NO: 12466.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX  
XX Homo sapiens.  
OS  
PN MO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483446/52.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -

XX Example 4; SEQ ID NO: 12466; 650bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.

XX Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;

Query Match 1.6%; Score 18; DB 22; Length 581;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 43 ctgtgctgctgctgtct 60  
|||  
Db 168 CTGTGCTGCTGCTGCT 151

## RESULT 11

AAK38184/C  
ID AAK38184 standard; DNA; 581 BP.

AC AAK38184;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 12741.

KM Human: bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 12741; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is one of  
XX the probes of the invention.

XX Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;

Query Match 1.6%; Score 18; DB 22; Length 581;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 ctgtgctgctgctgtct 60  
|||  
Db 168 CTGTGCTGCTGCTGCT 151

## RESULT 12

AA118978/C  
ID AA118978 standard; DNA; 581 BP.

AC AA118978;

DT 12-OCT-2001 (first entry)

DE Probe #8911 for gene expression analysis in human cervical cell sample.

KM Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Claim 25; SEQ ID NO 8911; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;

Query Match 1.6%; Score 18; DB 22; Length 581;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 ctgtgctgctgctgtct 60  
|||  
Db 168 CTGTGCTGCTGCTGCT 151

## RESULT 13

ID	AAI44114/C
XX	AAI44114 standard; DNA; 581 BP.
AC	AAI44114;
XX	
DT	17-OCT-2001 (first entry)
XX	
DE	Probe #12800 used to measure gene expression in human placenta sample.
XX	
KW	Probe; microarray; human; placenta; antenatal diagnosis;
KM	genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	MO200157272-A2.
PD	09-AUG-2001.
XX	
PE	30-JAN-2001; 2001WO-US00663.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-488897/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
XX	
PS	Claim 25; SEQ ID No 12800; 654bp; English.
XX	
CC	The present invention relates to single exon nucleic acid probes (SENP).
CC	The present sequence is one such probe. The probes are useful for
CC	producing a microarray for predicting, measuring and displaying gene
CC	expression in samples derived from human placenta. The probes are useful
CC	for antenatal diagnosis of human genetic disorders.
XX	
SQ	Sequence 581 BP; 115 A; 144 C; 201 G; 121 T; 0 other;
Query Match                      1.6%; Score 18; DB 22; Length 581;	
Best Local Similarity        100.0%; Pred. No. 68;	
Matches     18; Conservative     0; Mismatches     0; Indels     0; Gaps     0.	
QY	43 ctgtctgcgctactgtct 60      t   t   t   t    168 CTGTCGTCTGCTCTCTCT 151
RESULT     14	
ID	AAFI2222
XX	AAFI2222 standard; cDNA; 694 BP.
AC	AAFI2222;
XX	
DT	13-MAR-2001 (first entry)
XX	
DE	Aspergillus oryzae EST SEQ ID NO:4745.
XX	
KW	Multiple gene expression; filamentous fungal cell; EST;
KM	expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KM	Aspergillus oryzae; Trichoderma reesei; Identification; recombination;
KM	culture condition; environmental stress; spore morphogenesis;
KM	metabolic pathway engineering; catalytic pathway engineering; ss.

KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;  
 XX liver cancer; lung cancer; cytostatic; ss.  
 OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	9..1388
FT	/*tag- a
FT	/product- "STMP2"
FT	/transl_except- (pos:372..374,aa:Xaa)
FT	/note- "Xaa- Ala or Thr"
FT	/transl_except- (pos:999..1001,aa:Xaa)
FT	/note- "Xaa- Pro or Leu"

WO200172962-A2.

04-OCT-2001.

23-MAR-2001; 2001WO-US09410.

24-MAR-2000; 2000US-191929P.

(SAAT/) SAATCIOGLU F.

Saatcioglu F;

WPI; 2001-662926/76.

P-PSDB; AAU10190.

New polynucleotide for the diagnosis, prevention and treatment for prostate and testis disorders, particularly prostate cancer, comprises prostate-specific or testis-specific nucleic acids -

Claim 4; Fig 4L; 114pp; English..

CC The invention relates to substantially pure prostate-specific or  
 CC testis-specific polypeptides and the nucleic acids encoding them.  
 CC Also included are vectors and host cells expressing the proteins, a  
 CC transgenic animal expressing the protein, antibodies against the  
 CC proteins, probes for detecting the nucleic acids, antisense molecules  
 CC for the nucleic acids and methods of isolating modulators of the  
 CC proteins. Compounds that modulate the prostate specific or testis  
 CC specific polypeptide are useful to diagnose, prevent or treat disorders  
 CC of the testis or prostate particularly prostate cancer, benign  
 CC prostatic hyperplasia, acute prostatitis, testicular cancer,  
 CC cryptorchidism, undescended, retractile, ascending or vanished  
 CC testis. Other proliferative disorders for which the modulators may be  
 CC used include lymphoma, leukaemia, melanoma, ovarian cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and lung cancer. The  
 CC present sequence encodes a prostate specific protein, Six-transmembrane  
 CC protein of prostate 2, STMP2.

Sequence 1396 BP; 370 A; 312 C; 301 G; 413 T; 0 other;

Query Match

Best Local Similarity 1.6%; Score 18; DB 22; Length 1396;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY	33 tatttgctgctgctgctg 50
DB	621 tatttgctgctgctgctg 638

Search completed: May 7, 2002, 18:15:25  
 Job time: 5247 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 15:55:54 ; Search time 72.68 Seconds

(Without alignments)  
3842.671 Million cell updates/sec

Title: US-09-579-383-1

Perfect score: 1137  
Sequence: 1 atgaatttaccgtaaataa.....tcgttgaaagctcactaa 1137

Scoring table: OLIGO\_NTUC  
Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents.NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCrUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	18	1.6	3416 2	US-08-451-822A-15 Sequence 15, Appl
C 2	18	1.6	3416 4	US-08-323-430-15 Sequence 3, Appl
C 3	18	1.6	6614 4	US-09-150-460B-3 Sequence 15, Appl
C 4	17	1.5	969 2	US-08-932-978-1 Sequence 1, Appl
C 5	17	1.5	2976 3	US-08-714-918-65 Sequence 65, Appl
C 6	17	1.5	2976 4	US-09-265-315-65 Sequence 65, Appl
C 7	17	1.5	2976 4	US-09-265-315-65 Sequence 65, Appl
C 8	17	1.5	2976 4	US-09-266-417-65 Sequence 65, Appl
C 9	17	1.5	6638 2	US-08-070-301-2 Sequence 2, Appl
C 10	17	1.5	40352 3	US-08-846-111D-15 Sequence 17, Appl
C 11	17	1.5	48974 4	US-08-920-422-17 Sequence 17, Appl
C 12	16	1.4	723 4	US-09-651-656-2 Sequence 2, Appl
C 13	16	1.4	1001 2	US-08-982-956-2 Sequence 2, Appl
C 14	16	1.4	1001 3	US-09-228-317-2 Sequence 2, Appl
C 15	16	1.4	1256 1	US-08-176-427B-9 Sequence 9, Appl
C 16	16	1.4	1256 2	US-08-356-060A-5 Sequence 5, Appl
C 17	16	1.4	1256 4	US-08-460-900C-5 Sequence 5, Appl
C 18	16	1.4	1256 4	US-08-674-509B-5 Sequence 5, Appl
C 19	16	1.4	1256 4	US-08-954-698-5 Sequence 5, Appl
C 20	16	1.4	1317 3	US-08-886-886-10 Sequence 10, Appl
C 21	16	1.4	1405 1	US-08-390-162-3 Sequence 3, Appl
C 22	16	1.4	1405 1	US-08-685-945B-3 Sequence 3, Appl
C 23	16	1.4	1713 1	US-08-386-727-5 Sequence 5, Appl
C 24	16	1.4	1713 2	US-08-600-452A-5 Sequence 5, Appl
C 25	16	1.4	2107 1	US-08-390-162-1 Sequence 1, Appl
C 26	16	1.4	2107 1	US-08-685-945B-1 Sequence 1, Appl
C 27	16	1.4	2158 1	US-07-688-352C-3 Sequence 3, Appl

28	16	1.4	2158 2	US-08-474-379C-3 Sequence 3, Appl
29	16	1.4	2158 3	US-09-146-249A-3 Sequence 3, Appl
30	16	1.4	2158 3	US-08-206-168B-3 Sequence 3, Appl
31	16	1.4	2158 5	PCR-US91-02714-3 Sequence 3, Appl
32	16	1.4	2223 1	US-08-257-073-4 Sequence 4, Appl
33	16	1.4	2262 1	US-07-689-008-3 Sequence 3, Appl
34	16	1.4	2274 3	US-08-492-459-13 Sequence 13, Appl
35	16	1.4	2274 3	US-08-423-752-13 Sequence 13, Appl
36	16	1.4	2274 4	US-08-716-873-27 Sequence 27, Appl
37	16	1.4	2274 4	US-09-368-431-27 Sequence 27, Appl
38	16	1.4	2274 4	US-09-414-006-13 Sequence 13, Appl
C 39	16	1.4	2456 1	US-08-233-146-8 Sequence 8, Appl
C 40	16	1.4	2456 1	US-08-463-470-8 Sequence 8, Appl
C 41	16	1.4	2608 3	US-08-804-439A-12 Sequence 12, Appl
C 42	16	1.4	2608 3	US-08-720-229-12 Sequence 12, Appl
C 43	16	1.4	2639 4	US-09-374-135-1 Sequence 1, Appl
C 44	16	1.4	2805 4	US-09-651-656-4 Sequence 4, Appl
C 45	16	1.4	2947 1	US-08-457-176-1 Sequence 1, Appl
C 46	16	1.4	2947 1	US-08-457-175-1 Sequence 1, Appl
C 47	16	1.4	3096 5	PCT-US96-05320A-897 Sequence 897, App
C 48	16	1.4	3209 1	US-07-803-633A-12 Sequence 12, Appl
C 49	16	1.4	3355 2	US-08-933-821-3 Sequence 3, Appl
C 50	16	1.4	3355 3	US-08-960-507-3 Sequence 3, Appl
C 51	16	1.4	3355 4	US-09-136-828-3 Sequence 3, Appl
C 52	16	1.4	4409 4	US-09-653-403-2 Sequence 2, Appl
C 53	16	1.4	4431 4	US-09-532-803-8 Sequence 8, Appl
C 54	16	1.4	5829 4	US-09-004-838-109 Sequence 109, App
C 55	16	1.4	6048 4	US-09-634-920-3 Sequence 3, Appl
C 56	16	1.4	6911 1	US-08-311-174-4 Sequence 4, Appl
C 57	16	1.4	9540 1	US-07-689-008-1 Sequence 1, Appl
C 58	16	1.4	24417 2	US-08-846-762-1 Sequence 1, Appl
C 59	16	1.4	68750 4	US-09-335-409-1 Sequence 1, Appl
C 60	16	1.4	68750 4	US-09-568-102-1 Sequence 1, Appl
C 61	16	1.4	68750 4	US-09-567-969-1 Sequence 1, Appl
C 62	16	1.4	68750 4	US-09-568-480-1 Sequence 1, Appl
C 63	16	1.4	68750 4	US-09-568-486-1 Sequence 1, Appl
C 64	16	1.4	68750 4	US-09-568-472-1 Sequence 1, Appl
C 65	16	1.4	71989 4	US-09-443-501A-2 Sequence 2, Appl
C 66	16	1.4	152331 3	US-09-128-155-16 Sequence 16, Appl
C 67	16	1.4	176373 3	US-09-128-155-17 Sequence 17, Appl
C 68	15	1.3	49 6	5240845-49 Patent No. 5240845
C 69	15	1.3	53 6	5240845-48 Patent No. 5240845
C 70	15	1.3	159 4	US-08-836-500A-7 Sequence 7, Appl
C 71	15	1.3	216 4	US-08-836-500A-5 Sequence 5, Appl
C 72	15	1.3	257 1	US-08-741-406-3 Sequence 3, Appl
C 73	15	1.3	257 3	US-09-024-472-3 Sequence 3, Appl
C 74	15	1.3	263 3	US-08-881-771A-6 Sequence 6, Appl
C 75	15	1.3	268 3	US-08-928-881-20 Sequence 20, Appl
C 76	15	1.3	321 4	US-09-199-637A-434 Sequence 434, App
C 77	15	1.3	401 4	US-09-328-111-761 Sequence 761, App
C 78	15	1.3	402 2	US-08-193-078B-20 Sequence 20, Appl
C 79	15	1.3	406 1	US-08-471-780C-99 Sequence 99, Appl
C 80	15	1.3	406 1	US-08-467-282B-99 Sequence 99, Appl
C 81	15	1.3	406 2	US-08-471-282A-99 Sequence 99, Appl
C 82	15	1.3	406 2	US-08-466-710C-99 Sequence 99, Appl
C 83	15	1.3	406 3	US-08-471-284B-99 Sequence 99, Appl
C 84	15	1.3	406 3	US-08-468-739C-99 Sequence 99, Appl
C 85	15	1.3	406 3	US-08-990-823-47 Sequence 47, Appl
C 86	15	1.3	489 4	US-09-199-637A-430 Sequence 430, App
C 87	15	1.3	510 4	US-08-432-994A-7 Sequence 7, Appl
C 88	15	1.3	537 4	US-08-836-500A-3 Sequence 3, Appl
C 89	15	1.3	602 1	US-08-061-314A-1 Sequence 1, Appl
C 90	15	1.3	627 4	US-08-856-207A-212 Sequence 212, App
C 91	15	1.3	651 2	US-08-920-418-1 Sequence 1, Appl
C 92	15	1.3	651 4	US-09-238-479-1 Sequence 1, Appl
C 93	15	1.3	676 4	US-08-998-416-1142 Sequence 1142, Ap
C 94	15	1.3	683 4	US-08-456-200B-1 Sequence 1, Appl
C 95	15	1.3	699 4	US-08-476-102A-6 Sequence 6, Appl
C 96	15	1.3	709 3	US-08-961-083-95 Sequence 95, Appl
C 97	15	1.3	717 4	US-08-998-416-586 Patent No. 529273
C 98	15	1.3	720 6	529273-11 Patent No. 529273
C 99	15	1.3	730 2	US-08-743-637B-11 Sequence 11, Appl
C 100	15	1.3	730 3	US-08-526-840B-11 Sequence 11, Appl

101	15	1.3	770	4	US-09-328-111-646	Sequence 646, App	c 174	15	1.3	2153	4	US-08-984-709A-19	Sequence 19, Appl
c 102	15	1.3	778	4	US-09-207-223-7	Sequence 7, Appl1	c 175	15	1.3	2212	3	US-08-960-507-18	Sequence 18, Appl
c 103	15	1.3	800	4	US-09-207-223-6	Sequence 6, Appl1	c 176	15	1.3	2381	2	US-08-736-770-1	Sequence 4, Appl1
104	15	1.3	868	3	US-08-961-083-19	Sequence 19, Appl1	c 177	15	1.3	2420	5	PCT-US93-000227-1	Sequence 1, Appl1
c 105	15	1.3	906	1	US-08-100-874-1	Sequence 1, Appl1	c 178	15	1.3	2540	2	US-08-511-485-3	Sequence 3, Appl1
c 106	15	1.3	937	4	US-09-532-656-13	Sequence 13, Appl1	c 179	15	1.3	2540	3	US-09-392-580-1	Sequence 1, Appl1
c 107	15	1.3	951	4	US-09-189-637A-432	Sequence 432, App	c 180	15	1.3	2625	2	US-08-468-036-2	Sequence 2, Appl1
c 108	15	1.3	1007	4	US-08-836-500A-1	Sequence 1, Appl1	c 181	15	1.3	2625	2	US-08-376-434-2	Sequence 2, Appl1
c 109	15	1.3	1008	3	US-08-721-979A-13	Sequence 13, Appl	c 182	15	1.3	2691	3	US-09-212-971-9	Sequence 9, Appl1
c 110	15	1.3	1050	3	US-09-199-637A-428	Sequence 428, App	c 183	15	1.3	2691	3	US-08-800-929A-9	Sequence 9, Appl1
c 111	15	1.3	1072	5	US-08-096-181A-13	Sequence 13, Appl	c 184	15	1.3	2691	3	US-09-617-053A-9	Sequence 9, Appl1
c 112	15	1.3	1072	5	PCT-US94-08326-13	Sequence 13, Appl	c 185	15	1.3	2736	4	US-08-714-918-104	Sequence 104, App
c 113	15	1.3	1073	4	US-08-476-102A-1	Sequence 1, Appl1	c 186	15	1.3	2736	4	US-09-265-315-104	Sequence 104, App
c 114	15	1.3	1074	3	US-08-096-181A-11	Sequence 11, Appl	c 187	15	1.3	2736	4	US-09-265-315-104	Sequence 104, App
c 115	15	1.3	1074	5	PCT-US94-08326-11	Sequence 11, Appl	c 188	15	1.3	2736	4	US-09-265-315-104	Sequence 104, App
c 116	15	1.3	1077	4	US-08-945-515-1	Sequence 1, Appl1	c 189	15	1.3	2784	4	US-08-845-258-9	Sequence 9, Appl1
c 117	15	1.3	1096	3	US-08-881-771A-5	Sequence 5, Appl1	c 190	15	1.3	2784	4	US-08-950-571-9	Sequence 9, Appl1
c 118	15	1.3	1134	4	US-08-432-994A-9	Sequence 9, Appl1	c 191	15	1.3	2784	4	US-08-723-142A-9	Sequence 9, Appl1
c 119	15	1.3	1135	4	US-08-915-795-7	Sequence 7, Appl1	c 192	15	1.3	2802	1	US-08-215-805A-79	Sequence 79, Appl
c 120	15	1.3	1137	3	US-08-096-181A-9	Sequence 9, Appl1	c 193	15	1.3	2980	4	US-08-456-200B-12	Sequence 12, Appl
c 121	15	1.3	1137	5	PCT-US94-08326-9	Sequence 9, Appl1	c 194	15	1.3	3010	4	US-09-199-637A-168	Sequence 168, App
c 122	15	1.3	1207	3	US-08-872-979-4	Sequence 4, Appl1	c 195	15	1.3	3095	3	US-08-434-000A-7	Sequence 7, Appl1
c 123	15	1.3	1242	6	5240845-2	Patent No. 5240845	c 196	15	1.3	3095	4	US-09-312-157-7	Sequence 7, Appl1
c 124	15	1.3	1248	3	US-08-910-505-3	Sequence 3, Appl1	c 197	15	1.3	3172	1	US-08-314-309A-1	Sequence 1, Appl1
c 125	15	1.3	1251	3	US-08-910-505-1	Sequence 1, Appl1	c 198	15	1.3	3325	6	5240838-4	Patent No. 5240838
c 126	15	1.3	1262	6	5240845-3	Patent No. 5240845	c 199	15	1.3	3648	1	US-08-053-614-1	Sequence 1, Appl1
c 127	15	1.3	1335	4	US-08-915-795-6	Sequence 6, Appl1	c 200	15	1.3	3648	1	US-08-316-397B-1	Sequence 1, Appl1
c 128	15	1.3	1330	2	US-08-715-131-1	Sequence 1, Appl1	c 201	15	1.3	3648	2	US-09-034-306-1	Sequence 1, Appl1
c 129	15	1.3	1330	4	US-09-231-753-1	Sequence 1, Appl1	c 202	15	1.3	3648	2	US-09-259-437-1	Sequence 1, Appl1
c 130	15	1.3	1333	1	US-08-928-881-18	Sequence 18, Appl	c 203	15	1.3	3648	5	PCT-US93-08782-1	Sequence 1, Appl1
c 131	15	1.3	1333	1	US-07-602-824A-1	Sequence 1, Appl1	c 204	15	1.3	3848	5	US-08-215-805A-1	Sequence 1, Appl1
c 132	15	1.3	1333	1	US-07-963-451-1	Sequence 1, Appl1	c 205	15	1.3	3999	2	US-08-971-244-1	Sequence 1, Appl1
c 133	15	1.3	1393	1	US-08-261-577-6	Sequence 6, Appl1	c 206	15	1.3	3999	2	US-09-286-691-1	Sequence 1, Appl1
c 134	15	1.3	1420	2	US-08-909-965C-3	Sequence 3, Appl1	c 207	15	1.3	4220	4	US-09-183-846A-11	Sequence 11, Appl
c 135	15	1.3	1464	1	US-07-735-065-1	Sequence 1, Appl1	c 208	15	1.3	4465	1	US-08-180-195-1	Sequence 1, Appl1
c 136	15	1.3	1464	1	US-08-469-202-11	Sequence 11, Appl	c 209	15	1.3	4465	1	US-08-477-329-1	Sequence 1, Appl1
c 137	15	1.3	1464	2	US-08-484-434C-11	Sequence 11, Appl	c 210	15	1.3	4465	2	US-08-475-458-1	Sequence 1, Appl1
c 138	15	1.3	1476	3	US-08-713-118-5	Sequence 5, Appl1	c 211	15	1.3	4465	3	US-08-960-400-1	Sequence 1, Appl1
c 139	15	1.3	1476	3	US-09-452-007-5	Sequence 5, Appl1	c 212	15	1.3	4465	4	US-09-583-459A-1	Sequence 1, Appl1
c 140	15	1.3	1477	5	US-08-096-181A-7	Sequence 7, Appl1	c 213	15	1.3	4465	4	US-09-583-459A-1	Sequence 1, Appl1
c 141	15	1.3	1477	5	PCT-US94-08326-7	Sequence 7, Appl1	c 214	15	1.3	4465	4	US-09-583-459A-1	Sequence 1, Appl1
c 142	15	1.3	1485	4	US-08-936-165A-218	Sequence 218, App	c 215	15	1.3	4748	3	US-09-331-581-1	Sequence 1, Appl1
c 143	15	1.3	1500	1	US-08-592-214A-15	Sequence 15, Appl	c 216	15	1.3	4781	2	US-09-001-273-1	Sequence 1, Appl1
c 144	15	1.3	1500	3	US-08-659-188-15	Sequence 15, Appl	c 217	15	1.3	4781	4	US-08-843-459A-1	Sequence 1, Appl1
c 145	15	1.3	1500	3	US-08-655-227-15	Sequence 15, Appl	c 218	15	1.3	4821	1	US-08-053-614-3	Sequence 3, Appl1
c 146	15	1.3	1500	3	US-08-655-241-15	Sequence 15, Appl	c 219	15	1.3	4821	1	US-08-316-397B-3	Sequence 3, Appl1
c 147	15	1.3	1500	3	US-09-149-976-15	Sequence 15, Appl	c 220	15	1.3	4821	2	US-09-034-306-3	Sequence 3, Appl1
c 148	15	1.3	1500	4	US-09-398-326-15	Sequence 15, Appl	c 221	15	1.3	4821	2	US-09-259-437-3	Sequence 3, Appl1
c 149	15	1.3	1525	1	US-08-609-572-1	Sequence 1, Appl1	c 222	15	1.3	4821	5	PCT-US93-08782-3	Sequence 3, Appl1
c 150	15	1.3	1525	4	US-08-841-751-1	Sequence 1, Appl1	c 223	15	1.3	4847	3	US-09-061-400-1	Sequence 1, Appl1
c 151	15	1.3	1525	4	US-08-846-340-1	Sequence 1, Appl1	c 224	15	1.3	4937	2	US-08-622-166A-3	Sequence 3, Appl1
c 152	15	1.3	1525	4	US-08-846-344-1	Sequence 1, Appl1	c 225	15	1.3	5232	3	US-09-212-971-3	Sequence 3, Appl1
c 153	15	1.3	1591	1	US-08-728-956-3	Sequence 3, Appl1	c 226	15	1.3	5232	3	US-08-800-929A-3	Sequence 3, Appl1
c 154	15	1.3	1608	2	US-08-632-166A-1	Sequence 1, Appl1	c 227	15	1.3	5427	4	US-09-617-053A-3	Sequence 3, Appl1
c 155	15	1.3	1678	2	US-08-976-259-11	Sequence 11, Appl1	c 228	15	1.3	5427	2	US-08-168-917-1	Sequence 1, Appl1
c 156	15	1.3	1751	4	US-08-955-138-3	Sequence 3, Appl1	c 229	15	1.3	5427	2	US-08-460-510-1	Sequence 1, Appl1
c 157	15	1.3	1791	1	US-08-565-386-10	Sequence 10, Appl	c 230	15	1.3	5427	2	US-08-460-510-1	Sequence 1, Appl1
c 158	15	1.3	1830	1	US-08-343-733A-2	Sequence 2, Appl1	c 231	15	1.3	5427	2	US-08-460-510-1	Sequence 1, Appl1
c 159	15	1.3	1932	2	US-08-392-806A-1	Sequence 1, Appl1	c 232	15	1.3	5427	5	PCT-US92-00730-1	Sequence 1, Appl1
c 160	15	1.3	1932	4	US-09-257-490-1	Sequence 1, Appl1	c 233	15	1.3	5427	5	PCT-US92-00730-1	Sequence 1, Appl1
c 161	15	1.3	2026	2	US-08-149-097D-19	Sequence 19, Appl	c 234	15	1.3	5648	5	PCT-US93-00862-1	Sequence 1, Appl1
c 162	15	1.3	2038	2	US-08-736-770-2	Sequence 2, Appl1	c 235	15	1.3	5648	5	US-09-371-008-1	Sequence 1, Appl1
c 163	15	1.3	2100	2	US-08-511-485-9	Sequence 9, Appl1	c 236	15	1.3	5687	2	US-08-380-403A-3	Sequence 3, Appl1
c 164	15	1.3	2126	2	US-08-789-354-1	Sequence 1, Appl1	c 237	15	1.3	5904	5	US-08-309-512-1	Sequence 1, Appl1
c 165	15	1.3	2126	3	US-09-110-937-1	Sequence 1, Appl1	c 238	15	1.3	5904	5	PCT-US92-08756A-1	Sequence 1, Appl1
c 166	15	1.3	2126	3	US-09-058-725B-1	Sequence 1, Appl1	c 239	15	1.3	5931	3	US-08-783-774-1	Sequence 1, Appl1
c 167	15	1.3	2126	3	US-09-232-857-1	Sequence 1, Appl1	c 240	15	1.3	5962	1	US-08-188-582-10	Sequence 10, Appl
c 168	15	1.3	2144	2	US-08-149-097D-20	Sequence 20, Appl	c 241	15	1.3	5962	1	US-08-446-715-10	Sequence 10, Appl
c 169	15	1.3	2144	3	US-08-949-386-20	Sequence 20, Appl	c 242	15	1.3	6681	4	US-08-976-259-65	Sequence 65, Appl
c 170	15	1.3	2144	3	US-08-450-562-20	Sequence 20, Appl	c 243	15	1.3	7125	1	US-07-745-206A-1	Sequence 1, Appl1
c 171	15	1.3	2144	4	US-08-984-709A-20	Sequence 20, Appl	c 244	15	1.3	7125	2	US-08-311-563-1	Sequence 1, Appl1
c 172	15	1.3	2153	3	US-08-949-386-19	Sequence 19, Appl	c 245	15	1.3	7286	3	US-09-331-581-3	Sequence 3, Appl1
c 173	15	1.3	2153	3	US-08-450-562-19	Sequence 19, Appl	c 246	15	1.3	7635	1	US-08-453-543A-1	Sequence 1, Appl1

247	15	1.3	7635	1	US-08-455-543A-23	Sequence 23, Appl	C 320	14	1.2	89	3	US-08-722-719-15	Sequence 15, Appl
248	15	1.3	7635	2	US-08-193-078B-1	Sequence 1, Appl	C 321	14	1.2	163	1	US-08-248-474-25	Sequence 25, Appl
249	15	1.3	7635	2	US-08-193-078B-1	Sequence 29, Appl	C 322	14	1.2	163	3	US-08-756-849-25	Sequence 25, Appl
250	15	1.3	7635	2	US-08-223-305C-1	Sequence 1, Appl	C 323	14	1.2	176	4	US-08-991-789A-207	Sequence 207, App
251	15	1.3	7635	2	US-08-223-305C-23	Sequence 23, Appl	C 324	14	1.2	176	4	US-09-062-451-207	Sequence 207, App
252	15	1.3	7635	2	US-08-149-097D-1	Sequence 1, Appl	C 325	14	1.2	181	2	US-08-821-559A-6	Sequence 6, Appl
253	15	1.3	7635	3	US-08-949-386-1	Sequence 1, Appl	C 326	14	1.2	181	4	US-09-400-541-6	Sequence 40, Appl
254	15	1.3	7635	3	US-08-450-562-1	Sequence 1, Appl	C 327	14	1.2	249	1	US-07-872-644-40	Sequence 40, Appl
255	15	1.3	7635	3	US-08-984-709A-1	Sequence 1, Appl	C 328	14	1.2	249	1	US-08-297-494-40	Sequence 40, Appl
256	15	1.3	7938	3	US-09-331-581-14	Sequence 14, Appl	C 329	14	1.2	249	1	US-08-297-510-40	Sequence 40, Appl
257	15	1.3	8532	1	US-08-452-655B-1	Sequence 1, Appl	C 330	14	1.2	249	1	US-08-479-532-40	Sequence 40, Appl
258	15	1.3	8532	1	US-08-450-582-1	Sequence 1, Appl	C 331	14	1.2	249	1	US-08-435-526-40	Sequence 40, Appl
259	15	1.3	9606	3	US-07-741-940-1	Sequence 1, Appl	C 332	14	1.2	249	1	US-08-455-525-40	Sequence 40, Appl
260	15	1.3	9606	1	US-08-289-548A-1	Sequence 1, Appl	C 333	14	1.2	249	3	US-09-139-491-40	Sequence 40, Appl
261	15	1.3	9606	1	US-08-452-654-1	Sequence 1, Appl	C 334	14	1.2	249	5	PCT-US92-03222-10	Sequence 40, Appl
262	15	1.3	9606	2	US-08-370-235A-1	Sequence 1, Appl	C 335	14	1.2	278	4	US-09-060-756-3	Sequence 3, Appl
263	15	1.3	9717	4	US-09-251-645-1	Sequence 1, Appl	C 336	14	1.2	300	4	US-08-938-263-22	Sequence 22, Appl
264	15	1.3	10322	4	US-09-330-330-3	Sequence 3, Appl	C 337	14	1.2	309	4	US-09-221-298-4	Sequence 4, Appl
265	15	1.3	11459	4	US-09-454-721A-3	Sequence 3, Appl	C 338	14	1.2	313	4	US-08-976-259-110	Sequence 110, App
266	15	1.3	22671	4	US-08-976-259-14	Sequence 14, Appl	C 339	14	1.2	317	4	US-08-991-789A-181	Sequence 181, App
267	15	1.3	22846	2	US-08-469-461-3	Sequence 3, Appl	C 340	14	1.2	317	4	US-09-062-451-181	Sequence 181, App
268	15	1.3	22846	1	US-07-890-609-3	Sequence 3, Appl	C 341	14	1.2	320	4	US-09-481-161-1	Sequence 1, Appl
269	15	1.3	28958	1	US-08-258-261B-6	Sequence 6, Appl	C 342	14	1.2	321	4	US-09-087-465-21	Sequence 21, Appl
270	15	1.3	28958	1	US-08-456-837-6	Sequence 6, Appl	C 343	14	1.2	331	4	US-09-060-756-347	Sequence 347, App
271	15	1.3	28958	1	US-08-457-342-6	Sequence 6, Appl	C 344	14	1.2	344	2	US-08-875-972-17	Sequence 17, Appl
272	15	1.3	28958	1	US-08-458-676A-6	Sequence 6, Appl	C 345	14	1.2	386	4	US-09-060-756-364	Sequence 364, App
273	15	1.3	28958	1	US-08-458-076A-6	Sequence 6, Appl	C 346	14	1.2	389	4	US-09-481-161-3	Sequence 3, Appl
274	15	1.3	28958	1	US-08-764-233A-4	Sequence 4, Appl	C 347	14	1.2	404	1	US-08-222-616-21	Sequence 21, Appl
275	15	1.3	28958	1	US-08-457-335A-6	Sequence 6, Appl	C 348	14	1.2	404	5	PCT-US95-04228-21	Sequence 21, Appl
276	15	1.3	28958	1	US-08-729-214-6	Sequence 6, Appl	C 349	14	1.2	406	4	US-09-386-493-13	Sequence 13, Appl
277	15	1.3	28958	1	US-09-028-934-6	Sequence 6, Appl	C 350	14	1.2	423	1	US-08-470-179-135	Sequence 135, App
278	15	1.3	29793	4	US-09-302-812-38	Sequence 38, Appl	C 351	14	1.2	461	1	US-08-474-542A-278	Sequence 278, App
279	15	1.3	29793	4	US-09-511-477-38	Sequence 38, Appl	C 352	14	1.2	461	1	US-08-457-648-278	Sequence 278, App
280	15	1.3	35100	2	US-08-770-379-19	Sequence 19, Appl	C 353	14	1.2	490	5	PCT-US95-08295-23	Sequence 23, Appl
281	15	1.3	35100	4	US-08-757-669A-19	Sequence 19, Appl	C 354	14	1.2	537	2	US-08-675-508-3	Sequence 3, Appl
282	15	1.3	35100	4	US-09-230-371A-19	Sequence 19, Appl	C 355	14	1.2	545	4	US-09-227-357-149	Sequence 149, App
283	15	1.3	35524	3	US-08-923-137-1	Sequence 1, Appl	C 356	14	1.2	565	3	US-08-866-340-32	Sequence 32, Appl
284	15	1.3	37948	4	US-09-251-645-11	Sequence 11, Appl	C 357	14	1.2	565	4	US-09-103-875-38	Sequence 38, Appl
285	15	1.3	42235	4	US-09-199-637A-1	Sequence 1, Appl	C 358	14	1.2	578	2	US-08-619-542B-46	Sequence 46, Appl
286	15	1.3	43280	2	US-08-804-227C-1	Sequence 1, Appl	C 359	14	1.2	603	4	US-09-385-982-113	Sequence 113, App
287	15	1.3	49377	1	US-08-764-233A-1	Sequence 1, Appl	C 360	14	1.2	603	4	US-09-385-982-110	Sequence 110, App
288	15	1.3	59065	4	US-09-813-817-3	Sequence 3, Appl	C 361	14	1.2	617	4	US-09-328-111-577	Sequence 577, App
289	15	1.3	80161	4	US-09-036-987A-1	Sequence 1, Appl	C 362	14	1.2	617	4	US-09-385-982-222	Sequence 222, App
290	15	1.3	80161	4	US-09-370-700-1	Sequence 1, Appl	C 363	14	1.2	623	4	US-08-998-416-921	Sequence 921, App
291	15	1.3	87350	1	US-08-781-891-79	Sequence 79, Appl	C 364	14	1.2	627	3	US-09-181-183-9	Sequence 9, Appl
292	15	1.3	246240	2	US-08-724-394A-20	Sequence 20, Appl	C 365	14	1.2	627	4	US-09-277-700-9	Sequence 9, Appl
293	15	1.3	246240	2	US-08-724-394A-21	Sequence 21, Appl	C 366	14	1.2	629	2	US-08-670-186-1	Sequence 1, Appl
294	15	1.3	246240	2	US-08-724-394A-22	Sequence 22, Appl	C 367	14	1.2	630	4	US-08-646-665-13	Sequence 13, Appl
295	15	1.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl	C 368	14	1.2	630	4	US-08-646-665-13	Sequence 14, Appl
296	15	1.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl	C 369	14	1.2	630	5	PCT-US96-06053-13	Sequence 13, Appl
297	15	1.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl	C 370	14	1.2	630	5	PCT-US96-06053-14	Sequence 14, Appl
298	15	1.3	4411529	4	US-09-339-993-32	Sequence 32, Appl	C 371	14	1.2	643	4	US-08-936-165A-247	Sequence 247, App
299	14	1.2	18	3	US-09-559-303B-19	Sequence 19, Appl	C 372	14	1.2	643	4	US-08-861-774E-11	Sequence 31, Appl
300	14	1.2	24	4	US-09-175-828-19	Sequence 19, Appl	C 373	14	1.2	645	4	US-08-998-416-907	Sequence 907, App
301	14	1.2	24	4	US-07-989-160-10	Sequence 10, Appl	C 374	14	1.2	668	1	US-08-516-545-1	Sequence 1, Appl
302	14	1.2	25	1	US-08-182-961B-33	Sequence 33, Appl	C 375	14	1.2	677	4	US-09-328-111-281	Sequence 281, App
303	14	1.2	25	1	US-09-007-678B-33	Sequence 33, Appl	C 376	14	1.2	680	4	US-09-328-111-627	Sequence 627, App
304	14	1.2	25	4	US-08-021-667A-11	Sequence 11, Appl	C 377	14	1.2	689	1	US-08-106-507-9	Sequence 9, Appl
305	14	1.2	27	1	US-08-410-544-11	Sequence 11, Appl	C 378	14	1.2	689	1	US-08-446-922-7	Sequence 7, Appl
306	14	1.2	27	1	US-08-728-785A-11	Sequence 11, Appl	C 379	14	1.2	689	5	PCT-US93-10034-7	Sequence 7, Appl
307	14	1.2	27	1	US-08-089-910-1	Sequence 1, Appl	C 380	14	1.2	703	4	US-08-856-207A-21	Sequence 21, Appl
308	14	1.2	30	1	US-09-009-443-17	Sequence 1, Appl	C 381	14	1.2	713	2	US-08-580-545B-7	Sequence 7, Appl
309	14	1.2	31	4	US-08-087-772A-11	Sequence 11, Appl	C 382	14	1.2	713	4	US-09-262-653A-7	Sequence 7, Appl
310	14	1.2	35	2	US-08-585-684B-1528	Sequence 1528, Ap	C 383	14	1.2	714	4	US-08-971-090-2	Sequence 2, Appl
311	14	1.2	36	2	US-08-585-684B-1528	Sequence 1585, Ap	C 384	14	1.2	714	4	US-08-998-416-1009	Sequence 1009, App
312	14	1.2	36	3	US-08-722-719-12	Sequence 12, Appl	C 385	14	1.2	720	4	US-08-998-416-1009	Sequence 399, App
313	14	1.2	36	3	US-09-038-073-1528	Sequence 1528, Ap	C 386	14	1.2	766	4	US-08-836-236-3	Sequence 3, Appl
314	14	1.2	36	4	US-09-038-073-1585	Sequence 1585, Ap	C 387	14	1.2	785	3	US-08-969-644-15	Sequence 15, Appl
315	14	1.2	36	4	US-08-403-762A-65	Sequence 65, App	C 388	14	1.2	795	3	US-08-444-189-15	Sequence 15, Appl
316	14	1.2	37	1	US-08-952-793-353	Sequence 353, App	C 389	14	1.2	795	3	US-08-468-544-15	Sequence 15, Appl
317	14	1.2	61	4	PCT-US96-09455A-353	Sequence 353, App	C 390	14	1.2	807	3	US-08-871-483-10	Sequence 10, Appl
318	14	1.2	61	5			C 391	14	1.2	807	3	US-09-181-103-35	Sequence 35, Appl
319	14	1.2	61	5			C 392	14	1.2	816	3		

C 393	14	1.2	816	4	US-09-277-700-35	Sequence 35, Appl	466	14	1.2	1250	2	US-08-455-736-39	Sequence 39, Appl
394	14	1.2	830	4	US-08-998-416-419	Sequence 419, App	467	14	1.2	1250	2	US-08-971-217-39	Sequence 39, Appl
395	14	1.2	846	2	US-08-619-542B-43	Sequence 43, Appl	468	14	1.2	1250	4	US-09-350-600-39	Sequence 39, Appl
396	14	1.2	913	1	US-08-109-391A-1	Sequence 1, Appl	469	14	1.2	1251	4	US-09-242-095-1	Sequence 1, Appl
397	14	1.2	913	1	US-08-459-019A-1	Sequence 1, Appl	C 470	14	1.2	1272	1	US-08-343-682-2	Sequence 2, Appl
398	14	1.2	913	2	US-08-460-428A-1	Sequence 1, Appl	C 471	14	1.2	1299	2	US-08-904-278-5	Sequence 5, Appl
399	14	1.2	913	3	US-08-458-860A-1	Sequence 1, Appl	C 472	14	1.2	1299	4	US-09-222-594-5	Sequence 5, Appl
C 400	14	1.2	926	1	US-08-106-507-1	Sequence 1, Appl	C 473	14	1.2	1301	2	US-08-467-948A-7	Sequence 7, Appl
C 401	14	1.2	929	3	US-08-860-368B-4	Sequence 4, Appl	474	14	1.2	1301	3	US-08-467-947A-7	Sequence 7, Appl
C 402	14	1.2	930	4	US-09-227-357-61	Sequence 61, Appl	475	14	1.2	1303	3	US-08-700-186A-7	Sequence 1, Appl
C 403	14	1.2	930	4	US-09-227-357-146	Sequence 146, App	476	14	1.2	1303	2	US-08-914-981-1	Sequence 1, Appl
C 404	14	1.2	957	1	US-07-745-206A-16	Sequence 16, Appl	477	14	1.2	1303	3	US-09-116-115-1	Sequence 1, Appl
C 405	14	1.2	957	2	US-08-311-363-16	Sequence 16, Appl	478	14	1.2	1303	4	US-09-541-762-1	Sequence 1, Appl
C 406	14	1.2	998	4	US-09-122-400B-5	Sequence 5, Appl	479	14	1.2	1308	4	US-09-150-213-1	Sequence 1, Appl
C 407	14	1.2	998	4	US-09-132-400B-5	Sequence 5, Appl	480	14	1.2	1314	3	US-09-025-059-2	Sequence 2, Appl
C 408	14	1.2	1002	2	US-09-328-111-45	Sequence 445, App	481	14	1.2	1317	3	US-08-886-886-1	Sequence 1, Appl
C 409	14	1.2	1020	2	US-08-475-844-10	Sequence 10, Appl	482	14	1.2	1343	4	US-09-499-505-5	Sequence 5, Appl
C 410	14	1.2	1020	5	PCT-US95-08429-10	Sequence 10, Appl	483	14	1.2	1343	4	US-09-626-410-5	Sequence 5, Appl
C 411	14	1.2	1024	6	PCT-US95-16	Sequence 10, Appl	484	14	1.2	1343	4	US-09-116-188-5	Sequence 5, Appl
C 412	14	1.2	1027	4	US-09-276-531-24	Sequence 24, Appl	485	14	1.2	1343	4	US-09-626-410-5	Sequence 5, Appl
C 413	14	1.2	1032	4	US-09-146-319-1	Sequence 1, Appl	486	14	1.2	1343	4	US-09-626-343-5	Sequence 5, Appl
C 414	14	1.2	1032	4	US-09-175-973-1	Sequence 1, Appl	487	14	1.2	1345	2	US-08-702-153-3	Sequence 5, Appl
C 415	14	1.2	1038	4	US-09-004-838-128	Sequence 128, App	488	14	1.2	1356	4	US-09-371-913A-8	Sequence 3, Appl
C 416	14	1.2	1045	2	US-09-014-969-6	Sequence 1, Appl	489	14	1.2	1379	4	US-09-499-505-6	Sequence 6, Appl
C 417	14	1.2	1066	1	US-08-154-916-1	Sequence 1, Appl	490	14	1.2	1379	4	US-09-626-410-6	Sequence 6, Appl
C 418	14	1.2	1074	1	US-08-045-269C-3	Sequence 3, Appl	491	14	1.2	1379	4	US-09-116-188-6	Sequence 6, Appl
C 419	14	1.2	1074	3	US-08-371-680-3	Sequence 3, Appl	492	14	1.2	1379	4	US-09-626-047-6	Sequence 6, Appl
C 420	14	1.2	1074	5	PCT-US94-01198-3	Sequence 3, Appl	493	14	1.2	1379	4	US-09-626-343-6	Sequence 6, Appl
C 421	14	1.2	1075	3	US-09-116-115-16	Sequence 16, Appl	494	14	1.2	1380	4	US-09-499-505-4	Sequence 4, Appl
C 422	14	1.2	1075	4	US-09-341-762-16	Sequence 16, Appl	495	14	1.2	1380	4	US-09-626-410-4	Sequence 4, Appl
C 423	14	1.2	1094	4	US-09-144-918-3	Sequence 3, Appl	496	14	1.2	1380	4	US-09-116-188-4	Sequence 4, Appl
C 424	14	1.2	1095	2	US-09-139-424-1	Sequence 1, Appl	497	14	1.2	1380	4	US-09-626-047-4	Sequence 4, Appl
C 425	14	1.2	1096	1	US-08-525-507-4	Sequence 1, Appl	498	14	1.2	1380	4	US-09-626-343-4	Sequence 4, Appl
C 426	14	1.2	1121	1	US-08-207-412B-1	Sequence 1, Appl	499	14	1.2	1382	4	US-09-499-505-2	Sequence 2, Appl
C 427	14	1.2	1131	5	PCT-US95-02950-1	Sequence 1, Appl	500	14	1.2	1382	4	US-09-626-410-2	Sequence 2, Appl
C 428	14	1.2	1133	4	US-08-858-207A-3	Sequence 3, Appl	501	14	1.2	1382	4	US-09-116-188-2	Sequence 2, Appl
C 429	14	1.2	1140	2	US-08-904-278-7	Sequence 7, Appl	502	14	1.2	1382	4	US-09-626-047-2	Sequence 2, Appl
C 430	14	1.2	1140	3	US-09-222-594-7	Sequence 7, Appl	503	14	1.2	1384	3	US-09-626-343-2	Sequence 3, Appl
C 431	14	1.2	1158	3	US-08-860-368B-1	Sequence 1, Appl	C 504	14	1.2	1384	3	US-09-058-489-30	Sequence 30, Appl
C 432	14	1.2	1153	3	US-08-746-397-1	Sequence 1, Appl	505	14	1.2	1391	2	US-08-743-637B-7	Sequence 7, Appl
C 433	14	1.2	1157	1	US-07-960-985-1	Sequence 1, Appl	506	14	1.2	1391	2	US-08-526-840B-7	Sequence 7, Appl
C 434	14	1.2	1157	2	US-08-496-671-1	Sequence 1, Appl	507	14	1.2	1396	4	US-09-499-505-13	Sequence 13, Appl
C 435	14	1.2	1173	3	US-08-740-644-1	Sequence 1, Appl	508	14	1.2	1398	4	US-09-626-410-13	Sequence 13, Appl
C 436	14	1.2	1181	1	US-08-181-271A-27	Sequence 27, Appl	509	14	1.2	1398	4	US-09-116-188-13	Sequence 13, Appl
C 437	14	1.2	1181	1	US-08-449-315-27	Sequence 27, Appl	510	14	1.2	1398	4	US-09-626-047-13	Sequence 13, Appl
C 438	14	1.2	1181	1	US-08-444-803-27	Sequence 27, Appl	511	14	1.2	1398	4	US-09-626-343-13	Sequence 13, Appl
C 439	14	1.2	1181	1	US-08-449-043-27	Sequence 27, Appl	512	14	1.2	1400	4	US-09-144-918-1	Sequence 1, Appl
C 440	14	1.2	1181	1	US-08-456-265A-27	Sequence 27, Appl	513	14	1.2	1410	2	US-08-975-316-86	Sequence 86, Appl
C 441	14	1.2	1181	1	US-08-455-416-27	Sequence 27, Appl	C 514	14	1.2	1418	2	US-08-922-171-1	Sequence 1, Appl
C 442	14	1.2	1181	1	US-08-455-244-27	Sequence 27, Appl	515	14	1.2	1422	1	US-08-439-725A-1	Sequence 1, Appl
C 443	14	1.2	1181	1	US-08-454-876-27	Sequence 27, Appl	516	14	1.2	1422	2	US-08-867-471-1	Sequence 1, Appl
C 444	14	1.2	1181	2	US-08-457-364-27	Sequence 27, Appl	517	14	1.2	1422	3	US-08-705-245-18	Sequence 18, Appl
C 445	14	1.2	1181	2	US-08-456-262-27	Sequence 27, Appl	518	14	1.2	1422	5	PCT-US96-06664-1	Sequence 1, Appl
C 446	14	1.2	1181	2	US-08-456-340-27	Sequence 27, Appl	C 519	14	1.2	1428	2	US-08-904-278-3	Sequence 3, Appl
C 447	14	1.2	1181	2	US-08-455-736-27	Sequence 27, Appl	C 520	14	1.2	1428	4	US-09-222-594-3	Sequence 3, Appl
C 448	14	1.2	1181	2	US-08-971-217-27	Sequence 27, Appl	521	14	1.2	1430	4	US-09-499-505-3	Sequence 3, Appl
C 449	14	1.2	1181	2	US-09-350-600-27	Sequence 27, Appl	C 522	14	1.2	1430	4	US-09-357-251-17	Sequence 17, Appl
C 450	14	1.2	1197	3	US-08-935-263-11	Sequence 11, Appl	C 523	14	1.2	1430	4	US-09-626-410-3	Sequence 3, Appl
C 451	14	1.2	1206	4	US-08-861-774E-13	Sequence 13, Appl	524	14	1.2	1430	4	US-09-116-188-3	Sequence 3, Appl
C 452	14	1.2	1231	3	US-08-965-600-2	Sequence 2, Appl	525	14	1.2	1430	4	US-09-626-047-3	Sequence 3, Appl
C 453	14	1.2	1234	4	US-09-574-942-1	Sequence 1, Appl	526	14	1.2	1430	4	US-09-626-343-3	Sequence 3, Appl
C 454	14	1.2	1237	4	US-09-007-678B-48	Sequence 48, Appl	C 527	14	1.2	1430	2	US-08-692-787-6	Sequence 6, Appl
C 455	14	1.2	1250	1	US-08-181-271A-39	Sequence 39, Appl	C 528	14	1.2	1478	1	US-08-700-359-3	Sequence 3, Appl
C 456	14	1.2	1250	1	US-08-449-315-39	Sequence 39, Appl	529	14	1.2	1485	4	US-09-499-505-1	Sequence 1, Appl
C 457	14	1.2	1250	1	US-08-444-803-39	Sequence 39, Appl	530	14	1.2	1485	4	US-09-626-410-1	Sequence 1, Appl
C 458	14	1.2	1250	1	US-08-449-043-39	Sequence 39, Appl	531	14	1.2	1485	4	US-09-116-188-1	Sequence 1, Appl
C 459	14	1.2	1250	1	US-08-456-265A-39	Sequence 39, Appl	532	14	1.2	1485	4	US-09-626-047-1	Sequence 1, Appl
C 460	14	1.2	1250	1	US-08-455-416-39	Sequence 39, Appl	533	14	1.2	1485	4	US-09-626-343-1	Sequence 1, Appl
C 461	14	1.2	1250	1	US-08-455-244-39	Sequence 39, Appl	534	14	1.2	1516	3	US-09-626-410-1	Sequence 1, Appl
C 462	14	1.2	1250	1	US-08-454-876-39	Sequence 39, Appl	535	14	1.2	1516	3	US-09-344-914-1	Sequence 1, Appl
C 463	14	1.2	1250	2	US-08-457-364-39	Sequence 39, Appl	536	14	1.2	1558	4	US-08-336-165A-200	Sequence 200, App
C 464	14	1.2	1250	2	US-08-456-262-39	Sequence 39, Appl	537	14	1.2	1545	1	US-07-621-670-1	Sequence 1, Appl
C 465	14	1.2	1250	2	US-08-456-240-39	Sequence 39, Appl	538	14	1.2	1546	3	US-08-961-083-21	Sequence 21, Appl

539	14	1.2	1552	2	US-08-752-760A-3	Sequence 3, Appli	612	14	1.2	2294	2	US-08-676-279-49	Sequence 49, Appli
540	14	1.2	1553	3	US-08-492-459-21	Sequence 21, Appli	C 613	14	1.2	2303	2	US-08-480-229C-9	Sequence 9, Appli
541	14	1.2	1553	3	US-08-423-752-21	Sequence 21, Appli	C 614	14	1.2	2303	2	US-08-659-235C-9	Sequence 9, Appli
542	14	1.2	1553	4	US-08-716-873-35	Sequence 35, Appli	C 615	14	1.2	2308	2	US-08-480-229C-28	Sequence 28, Appli
543	14	1.2	1553	4	US-09-368-431-35	Sequence 35, Appli	C 616	14	1.2	2308	2	US-08-659-235C-28	Sequence 28, Appli
544	14	1.2	1553	4	US-09-414-006-21	Sequence 21, Appli	C 617	14	1.2	2338	2	US-08-425-069-1	Sequence 1, Appli
545	14	1.2	1554	3	US-08-045-269C-1	Sequence 1, Appli	C 618	14	1.2	2338	2	US-08-317-844B-1	Sequence 1, Appli
546	14	1.2	1554	3	US-08-371-680-1	Sequence 1, Appli	C 619	14	1.2	2339	5	PCT-US92-09325-3	Sequence 3, Appli
547	14	1.2	1554	5	PCT-US94-01198-1	Sequence 1, Appli	C 620	14	1.2	2340	5	PCT-US92-09325-1	Sequence 1, Appli
548	14	1.2	1558	5	PCT-US95-11808-2	Sequence 2, Appli	C 621	14	1.2	2349	2	US-08-184-009-145	Sequence 145, App
549	14	1.2	1571	1	US-08-726-525-1	Sequence 1, Appli	C 622	14	1.2	2349	2	US-08-458-356-145	Sequence 145, App
550	14	1.2	1571	1	US-08-487-942-1	Sequence 1, Appli	C 623	14	1.2	2349	4	US-08-460-736-145	Sequence 145, App
551	14	1.2	1571	2	US-08-726-036A-1	Sequence 1, Appli	C 624	14	1.2	2383	4	US-08-232-358-18	Sequence 18, App
552	14	1.2	1571	4	US-09-083-516-1	Sequence 1, Appli	C 625	14	1.2	2383	4	US-08-786-164-18	Sequence 18, Appli
553	14	1.2	1573	3	US-09-249-336-1	Sequence 1, Appli	C 626	14	1.2	2413	4	US-09-518-046-1	Sequence 1, Appli
554	14	1.2	1573	4	US-09-249-338-1	Sequence 1, Appli	C 627	14	1.2	2416	4	US-09-261-416-1	Sequence 1, Appli
555	14	1.2	1584	4	US-09-318-448-31	Sequence 31, Appli	C 628	14	1.2	2434	2	US-08-184-009-144	Sequence 144, App
C 556	14	1.2	1593	2	US-08-524-828-2	Sequence 2, Appli	C 629	14	1.2	2434	2	US-08-485-356-144	Sequence 144, App
C 557	14	1.2	1593	2	US-08-975-114A-2	Sequence 2, Appli	C 630	14	1.2	2434	4	US-08-460-736-144	Sequence 144, App
C 558	14	1.2	1593	2	US-08-849-281A-2	Sequence 2, Appli	C 631	14	1.2	2457	2	US-08-486-013-68	Sequence 68, Appli
C 559	14	1.2	1601	3	US-08-735-491-1	Sequence 1, Appli	C 632	14	1.2	2457	2	US-08-482-279-68	Sequence 68, Appli
C 560	14	1.2	1608	2	US-08-855-518-2	Sequence 2, Appli	C 633	14	1.2	2457	2	US-08-342-268-68	Sequence 68, Appli
C 561	14	1.2	1617	2	US-08-540-118-2	Sequence 2, Appli	C 634	14	1.2	2457	3	US-09-015-968-68	Sequence 68, Appli
C 562	14	1.2	1649	1	US-07-688-352C-35	Sequence 35, Appli	C 635	14	1.2	2457	4	US-09-397-386-68	Sequence 68, Appli
563	14	1.2	1649	2	US-08-474-379C-35	Sequence 35, Appli	C 636	14	1.2	2485	4	US-08-637-823B-1	Sequence 1, Appli
564	14	1.2	1649	3	US-09-146-249A-35	Sequence 35, Appli	C 637	14	1.2	2493	3	US-08-945-056-3	Sequence 3, Appli
565	14	1.2	1649	3	US-08-206-188B-35	Sequence 35, Appli	C 638	14	1.2	2544	3	US-09-518-046-3	Sequence 3, Appli
566	14	1.2	1649	5	PCT-US91-02714-34	Sequence 34, Appli	C 639	14	1.2	2551	2	US-08-486-013-70	Sequence 70, Appli
567	14	1.2	1664	3	US-09-339-993-1	Sequence 1, Appli	C 640	14	1.2	2551	2	US-08-482-279-70	Sequence 70, Appli
C 568	14	1.2	1697	1	US-08-343-027A-1	Sequence 1, Appli	C 641	14	1.2	2551	2	US-08-342-268-70	Sequence 70, Appli
C 569	14	1.2	1703	4	US-09-370-253-9	Sequence 9, Appli	C 642	14	1.2	2551	4	US-09-015-968-70	Sequence 70, Appli
570	14	1.2	1725	2	US-08-907-166-11	Sequence 11, Appli	C 643	14	1.2	2551	4	US-09-397-386-70	Sequence 70, Appli
571	14	1.2	1738	2	US-08-932-376A-1	Sequence 1, Appli	C 644	14	1.2	2577	4	US-08-628-434-1	Sequence 1, Appli
C 572	14	1.2	1746	3	US-08-951-130-5	Sequence 5, Appli	645	14	1.2	2600	4	US-08-628-434-1	Sequence 1, Appli
C 573	14	1.2	1764	2	US-08-918-723-2	Sequence 2, Appli	646	14	1.2	2607	1	US-08-369-796-5	Sequence 5, Appli
C 574	14	1.2	1764	2	US-09-237-507-2	Sequence 2, Appli	647	14	1.2	2607	5	PCT-US95-17025-5	Sequence 5, Appli
C 575	14	1.2	1767	3	US-08-930-894-1	Sequence 1, Appli	648	14	1.2	2607	2	US-08-852-091-5	Sequence 5, Appli
576	14	1.2	1788	1	US-08-035-928-1	Sequence 1, Appli	649	14	1.2	2607	3	US-08-820-754-5	Sequence 5, Appli
C 577	14	1.2	1789	1	US-08-455-543A-29	Sequence 1, Appli	650	14	1.2	2607	3	US-08-956-862-5	Sequence 5, Appli
C 578	14	1.2	1789	2	US-08-223-305C-29	Sequence 29, Appli	651	14	1.2	2607	3	US-08-956-869-5	Sequence 5, Appli
C 579	14	1.2	1848	2	US-08-622-740-7	Sequence 7, Appli	652	14	1.2	2607	4	US-08-948-547-7	Sequence 5, Appli
580	14	1.2	1848	3	US-08-440-689-7	Sequence 7, Appli	653	14	1.2	2607	5	US-08-956-863A-5	Sequence 5, Appli
581	14	1.2	1848	4	US-09-122-399-7	Sequence 7, Appli	654	14	1.2	2607	5	PCT-US95-17025-5	Sequence 5, Appli
582	14	1.2	1904	2	US-08-933-750C-95	Sequence 95, Appli	655	14	1.2	2657	1	US-07-718-355-2	Sequence 2, Appli
583	14	1.2	1904	3	US-09-234-613-95	Sequence 95, Appli	C 656	14	1.2	2657	1	US-08-161-999-2	Sequence 2, Appli
C 584	14	1.2	1921	3	US-08-482-677-3	Sequence 3, Appli	C 657	14	1.2	2671	5	PCT-US95-03552-1	Sequence 1, Appli
C 585	14	1.2	1926	1	US-08-152-019A-43	Sequence 43, Appli	C 658	14	1.2	2682	2	US-08-867-941-7	Sequence 7, Appli
586	14	1.2	1950	2	US-08-472-659-30	Sequence 30, Appli	C 659	14	1.2	2682	2	US-08-867-941-7	Sequence 7, Appli
587	14	1.2	1950	2	US-08-474-661-30	Sequence 30, Appli	C 660	14	1.2	2682	4	US-09-074-658-7	Sequence 7, Appli
588	14	1.2	1950	2	US-08-611-977-30	Sequence 30, Appli	C 661	14	1.2	2682	4	US-09-074-658-7	Sequence 7, Appli
589	14	1.2	1964	1	US-08-132-168A-31	Sequence 31, Appli	662	14	1.2	2694	2	US-08-867-941-2	Sequence 2, Appli
590	14	1.2	1964	4	US-08-687-590-57	Sequence 57, Appli	C 663	14	1.2	2694	4	US-08-867-941-2	Sequence 2, Appli
591	14	1.2	2019	3	US-08-926-636-1	Sequence 1, Appli	664	14	1.2	2694	4	US-09-074-658-2	Sequence 2, Appli
592	14	1.2	2019	4	US-09-282-305-5	Sequence 5, Appli	C 665	14	1.2	2694	4	US-09-074-658-2	Sequence 2, Appli
593	14	1.2	2031	1	US-08-217-299-2	Sequence 2, Appli	C 666	14	1.2	2700	3	US-09-022-875-1	Sequence 1, Appli
C 594	14	1.2	2074	4	US-08-630-915A-19	Sequence 19, Appli	C 667	14	1.2	2746	2	US-09-016-000-12	Sequence 12, Appli
595	14	1.2	2097	2	US-08-602-725-35	Sequence 35, Appli	C 668	14	1.2	2756	1	US-08-187-793-3	Sequence 3, Appli
596	14	1.2	2167	4	US-09-489-869-10	Sequence 10, Appli	C 669	14	1.2	2784	1	US-08-104-073-3	Sequence 3, Appli
597	14	1.2	2220	1	US-08-389-459A-16	Sequence 16, Appli	C 670	14	1.2	2784	2	US-08-351-413-10	Sequence 10, Appli
598	14	1.2	2220	3	US-08-932-376A-3	Sequence 3, Appli	C 671	14	1.2	2784	2	US-09-025-583-10	Sequence 10, Appli
599	14	1.2	2220	3	US-08-967-867A-16	Sequence 16, Appli	C 672	14	1.2	2799	3	US-08-968-752B-5	Sequence 5, Appli
C 600	14	1.2	2238	4	US-08-939-366-5	Sequence 5, Appli	673	14	1.2	2839	3	US-08-468-856B-5	Sequence 5, Appli
C 601	14	1.2	2242	3	US-09-400-742-1	Sequence 1, Appli	674	14	1.2	2839	3	US-08-468-859A-5	Sequence 5, Appli
C 602	14	1.2	2242	4	US-08-618-651A-1	Sequence 1, Appli	675	14	1.2	3046	1	US-08-525-507-7	Sequence 7, Appli
C 603	14	1.2	2242	4	US-09-215-252-1	Sequence 1, Appli	676	14	1.2	3133	4	US-08-103-069-1	Sequence 1, Appli
C 604	14	1.2	2247	2	US-08-524-828-1	Sequence 1, Appli	C 677	14	1.2	3164	1	US-08-188-228-19	Sequence 19, Appli
C 605	14	1.2	2247	2	US-08-975-114A-1	Sequence 1, Appli	C 678	14	1.2	3164	1	US-08-332-643-43	Sequence 43, Appli
C 606	14	1.2	2264	1	US-08-232-538-16	Sequence 16, Appli	C 679	14	1.2	3164	1	US-08-332-638-49	Sequence 49, Appli
C 607	14	1.2	2264	4	US-08-786-164-16	Sequence 16, Appli	680	14	1.2	3166	2	US-07-688-352C-23	Sequence 23, Appli
C 608	14	1.2	2276	4	US-08-637-823B-24	Sequence 24, Appli	681	14	1.2	3166	2	US-08-474-379C-23	Sequence 23, Appli
609	14	1.2	2276	4	US-07-861-458C-38	Sequence 38, Appli	682	14	1.2	3166	3	US-09-146-249A-23	Sequence 23, Appli
610	14	1.2	2280	3	US-09-272-114A-2	Sequence 2, Appli	683	14	1.2	3186	3	US-08-206-188B-23	Sequence 23, Appli
C 611	14	1.2	2292	4	US-09-142-956B-1	Sequence 1, Appli	684	14	1.2	3186	5	PCT-US91-02714-22	Sequence 22, Appli

685	14	1.2	3205	4	US-09-651-656-102	Sequence 102, App	758	14	1.2	4068	3	US-09-146-249A-58	Sequence 58, Appl
686	14	1.2	3234	1	US-08-286-325A-7	Sequence 7, Appl	759	14	1.2	4068	3	US-08-206-188B-58	Sequence 58, Appl
C 687	14	1.2	3250	3	US-08-617-860B-1	Sequence 1, Appl	C 760	14	1.2	4071	4	US-09-098-107A-1	Sequence 1, Appl
C 688	14	1.2	3293	2	US-08-442-809A-75	Sequence 75, Appl	C 761	14	1.2	4071	4	US-09-483-539-1	Sequence 1, Appl
C 689	14	1.2	3295	4	US-09-336-447A-8	Sequence 8, Appl	C 762	14	1.2	4106	1	US-08-434-823-1	Sequence 1, Appl
C 690	14	1.2	3318	1	US-08-187-793-3	Sequence 3, Appl	C 763	14	1.2	4106	1	US-08-457-366-1	Sequence 1, Appl
C 691	14	1.2	3350	3	US-08-617-860B-3	Sequence 3, Appl	C 764	14	1.2	4131	1	US-07-872-444-38	Sequence 38, Appl
C 692	14	1.2	3358	3	US-09-248-571-2	Sequence 2, Appl	C 765	14	1.2	4131	1	US-08-297-494-38	Sequence 38, Appl
C 693	14	1.2	3393	4	US-09-104-324B-1	Sequence 1, Appl	C 766	14	1.2	4131	1	US-08-297-510-38	Sequence 38, Appl
C 694	14	1.2	3393	4	US-09-162-713-1	Sequence 1, Appl	C 767	14	1.2	4131	1	US-08-485-588-4	Sequence 4, Appl
C 695	14	1.2	3450	2	US-08-545-562A-6	Sequence 6, Appl	C 768	14	1.2	4131	1	US-08-484-565-4	Sequence 4, Appl
C 696	14	1.2	3456	3	US-09-027-166-10	Sequence 10, Appl	C 769	14	1.2	4131	1	US-08-479-532-38	Sequence 38, Appl
C 697	14	1.2	3494	4	US-08-976-259-84	Sequence 8, Appl	C 770	14	1.2	4131	1	US-08-455-526-38	Sequence 38, Appl
C 698	14	1.2	3496	4	US-09-097-767A-37	Sequence 37, Appl	C 771	14	1.2	4131	1	US-08-455-525-38	Sequence 38, Appl
C 699	14	1.2	3544	2	US-08-485-139-3	Sequence 3, Appl	C 772	14	1.2	4131	2	US-08-480-751-4	Sequence 4, Appl
C 700	14	1.2	3544	3	US-08-750-357-3	Sequence 3, Appl	C 773	14	1.2	4131	2	US-08-943-986-4	Sequence 4, Appl
C 701	14	1.2	3568	3	US-08-894-440-3	Sequence 3, Appl	C 774	14	1.2	4131	2	US-08-353-784-4	Sequence 4, Appl
C 702	14	1.2	3658	4	US-09-458-093-3	Sequence 3, Appl	C 775	14	1.2	4131	3	US-09-139-491-38	Sequence 38, Appl
C 703	14	1.2	3658	4	US-08-894-440-3	Sequence 3, Appl	C 776	14	1.2	4131	3	US-08-484-719B-4	Sequence 4, Appl
C 704	14	1.2	3704	2	US-08-702-153-1	Sequence 1, Appl	C 777	14	1.2	4131	4	US-08-484-159-4	Sequence 4, Appl
C 705	14	1.2	3751	2	US-08-609-230A-8	Sequence 42, Appl	C 778	14	1.2	4131	5	PCT-US92-03222-38	Sequence 38, Appl
C 706	14	1.2	3789	1	US-07-872-644-42	Sequence 42, Appl	C 779	14	1.2	4140	3	US-08-894-731-2	Sequence 2, Appl
C 707	14	1.2	3789	1	US-08-297-494-42	Sequence 42, Appl	C 780	14	1.2	4235	1	US-08-021-601-3	Sequence 3, Appl
C 708	14	1.2	3789	1	US-08-479-532-42	Sequence 42, Appl	C 781	14	1.2	4235	1	US-08-082-849B-3	Sequence 3, Appl
C 709	14	1.2	3789	1	US-08-455-526-42	Sequence 42, Appl	C 782	14	1.2	4235	5	PCT-US94-0162A-3	Sequence 3, Appl
C 710	14	1.2	3789	1	US-08-455-525-42	Sequence 42, Appl	C 783	14	1.2	4236	1	US-08-810-116-7	Sequence 7, Appl
C 711	14	1.2	3789	3	US-09-139-491-42	Sequence 42, Appl	C 784	14	1.2	4236	2	US-07-930-548A-7	Sequence 7, Appl
C 712	14	1.2	3789	5	PCT-US92-03222-42	Sequence 42, Appl	C 785	14	1.2	4284	4	US-08-525-507-14	Sequence 14, Appl
C 713	14	1.2	3796	2	US-08-762-308-11	Sequence 11, Appl	C 786	14	1.2	4286	4	US-09-413-304-7	Sequence 7, Appl
C 714	14	1.2	3813	3	US-08-650-000-3	Sequence 3, Appl	C 787	14	1.2	4291	3	US-08-417-210A-81	Sequence 81, Appl
C 715	14	1.2	3813	6	5395760-3	Patent No. 3395760	C 788	14	1.2	4337	2	US-09-187-049-1	Sequence 2, Appl
C 716	14	1.2	3814	4	US-09-302-812-5	Sequence 5, Appl	C 789	14	1.2	4337	3	US-08-559-038-72	Sequence 1, Appl
C 717	14	1.2	3814	4	US-09-511-477-5	Sequence 5, Appl	C 790	14	1.2	4437	4	US-09-175-828-72	Sequence 72, Appl
C 718	14	1.2	3820	3	US-08-990-140-3	Sequence 3, Appl	C 791	14	1.2	4465	1	US-08-180-195-1	Sequence 1, Appl
C 719	14	1.2	3820	3	US-09-546-238-3	Sequence 3, Appl	C 792	14	1.2	4465	1	US-08-477-329-1	Sequence 1, Appl
C 720	14	1.2	3827	1	US-08-170-294-6	Sequence 6, Appl	C 793	14	1.2	4465	2	US-08-475-458-1	Sequence 1, Appl
C 721	14	1.2	3827	2	US-08-664-855-6	Sequence 6, Appl	C 794	14	1.2	4465	3	US-08-980-400-1	Sequence 1, Appl
C 722	14	1.2	3827	2	US-08-718-751-1	Sequence 1, Appl	C 795	14	1.2	4465	4	US-09-583-459A-1	Sequence 1, Appl
C 723	14	1.2	3827	3	US-09-049-289-6	Sequence 6, Appl	C 796	14	1.2	4465	4	US-09-583-410-1	Sequence 1, Appl
C 724	14	1.2	3829	2	US-08-631-097-8	Sequence 8, Appl	C 797	14	1.2	4465	4	US-09-583-449A-1	Sequence 1, Appl
C 725	14	1.2	3829	4	US-08-810-712-6	Sequence 6, Appl	C 798	14	1.2	4490	4	US-09-476-366A-1	Sequence 1, Appl
C 726	14	1.2	3867	4	US-09-347-114A-81	Sequence 81, Appl	C 799	14	1.2	4592	1	US-08-472-934-9	Sequence 9, Appl
C 727	14	1.2	3880	2	US-08-942-521B-1	Sequence 1, Appl	C 800	14	1.2	4592	2	US-08-323-460A-9	Sequence 9, Appl
C 728	14	1.2	3880	3	US-09-192-702-1	Sequence 1, Appl	C 801	14	1.2	4592	2	US-08-461-146C-9	Sequence 9, Appl
C 729	14	1.2	3890	4	US-08-445-474-1	Sequence 1, Appl	C 802	14	1.2	4592	2	US-08-461-145C-9	Sequence 9, Appl
C 730	14	1.2	3890	5	PCT-US94-02612-1	Sequence 1, Appl	C 803	14	1.2	4629	2	US-08-484-891-7	Sequence 7, Appl
C 731	14	1.2	3913	3	US-08-472-934-7	Sequence 7, Appl	C 804	14	1.2	4656	4	US-09-425-665-1	Sequence 1, Appl
C 732	14	1.2	3913	2	US-08-323-460A-7	Sequence 7, Appl	C 805	14	1.2	4656	4	US-09-685-668-1	Sequence 1, Appl
C 733	14	1.2	3913	2	US-08-461-146C-7	Sequence 7, Appl	C 806	14	1.2	4670	3	US-08-717-294-41	Sequence 41, Appl
C 734	14	1.2	3913	3	US-08-461-145C-7	Sequence 7, Appl	C 807	14	1.2	4758	4	US-09-191-647-1	Sequence 1, Appl
C 735	14	1.2	3913	4	US-08-628-829-11	Sequence 11, Appl	C 808	14	1.2	4758	4	US-09-540-245A-1	Sequence 1, Appl
C 736	14	1.2	3943	2	US-08-369-796-3	Sequence 3, Appl	C 809	14	1.2	4758	4	US-09-540-153-1	Sequence 1, Appl
C 737	14	1.2	3943	2	US-08-852-091-3	Sequence 3, Appl	C 810	14	1.2	4771	3	US-08-840-062-3	Sequence 3, Appl
C 738	14	1.2	3943	2	US-08-820-754-3	Sequence 3, Appl	C 811	14	1.2	4800	5	PCT-US94-07779-1	Sequence 1, Appl
C 739	14	1.2	3943	3	US-08-856-652-3	Sequence 3, Appl	C 812	14	1.2	4800	5	US-08-867-030B-5	Sequence 5, Appl
C 740	14	1.2	3943	3	US-08-856-869-3	Sequence 3, Appl	C 813	14	1.2	4951	2	PCT-US95-06119-5	Sequence 5, Appl
C 741	14	1.2	3943	3	US-08-948-547-3	Sequence 3, Appl	C 814	14	1.2	4951	2	US-08-920-812-18	Sequence 18, Appl
C 742	14	1.2	3943	4	US-09-364-970-10	Sequence 10, Appl	C 815	14	1.2	4954	1	US-08-920-827-18	Sequence 18, Appl
C 743	14	1.2	3943	4	US-08-956-653A-3	Sequence 3, Appl	C 816	14	1.2	4954	1	US-08-921-177-18	Sequence 18, Appl
C 744	14	1.2	3943	5	PCT-US95-17025-3	Sequence 3, Appl	C 817	14	1.2	4954	1	US-08-362-577C-18	Sequence 18, Appl
C 745	14	1.2	3987	1	US-07-688-352C-19	Sequence 19, Appl	C 818	14	1.2	4954	2	US-08-920-828-18	Sequence 18, Appl
C 746	14	1.2	3987	2	US-08-474-379C-19	Sequence 19, Appl	C 819	14	1.2	4954	4	US-09-670-618-14	Sequence 14, Appl
C 747	14	1.2	3987	3	US-09-146-249A-19	Sequence 19, Appl	C 820	14	1.2	4954	4	US-09-364-862-14	Sequence 14, Appl
C 748	14	1.2	3987	5	US-08-206-188B-19	Sequence 19, Appl	C 821	14	1.2	5027	3	US-08-680-506-2	Sequence 2, Appl
C 749	14	1.2	3987	5	PCT-US91-02714-19	Sequence 19, Appl	C 822	14	1.2	5035	2	US-08-882-083-1	Sequence 1, Appl
C 750	14	1.2	4002	2	US-08-331-488A-1	Sequence 1, Appl	C 823	14	1.2	5035	2	US-08-858-107-1	Sequence 1, Appl
C 751	14	1.2	4002	2	US-08-996-545-1	Sequence 1, Appl	C 824	14	1.2	5035	2	US-09-243-539-1	Sequence 1, Appl
C 752	14	1.2	4002	2	US-08-996-545-1	Sequence 1, Appl	C 825	14	1.2	5236	5	PCT-US91-09422-16	Sequence 16, Appl
C 753	14	1.2	4002	4	US-09-328-320-1	Sequence 1, Appl	C 826	14	1.2	5362	2	US-08-853-310-3	Sequence 3, Appl
C 754	14	1.2	4002	4	US-09-328-320-1	Sequence 1, Appl	C 827	14	1.2	5399	1	US-08-368-071-9	Sequence 9, Appl
C 755	14	1.2	4003	4	US-09-087-465-1	Sequence 1, Appl	C 828	14	1.2	5399	5	US-08-458-181-9	Sequence 9, Appl
C 756	14	1.2	4042	4	US-08-406-030A-17	Sequence 17, Appl	C 829	14	1.2	5399	5	PCT-US93-02172-9	Sequence 9, Appl
C 757	14	1.2	4068	2	US-08-474-379C-58	Sequence 58, Appl	C 830	14	1.2	5414	4	US-08-628-829-13	Sequence 13, Appl

831	14	1.2	5427	1	US-08-168-917-1	Sequence 1, Appl 1	904	14	1.2	7317	3	US-08-470-297A-2	Sequence 2, Appl 1
832	14	1.2	5427	2	US-08-460-510-1	Sequence 1, Appl 1	905	14	1.2	7317	5	PCT-US91-07149-2	Sequence 2, Appl 1
833	14	1.2	5427	2	US-08-460-490-1	Sequence 1, Appl 1	906	14	1.2	7317	5	US-08-440-787A-2	Sequence 2, Appl 1
834	14	1.2	5427	3	US-08-462-728-3	Sequence 1, Appl 1	907	14	1.2	7320	5	US-08-367-685-2	Sequence 2, Appl 1
835	14	1.2	5427	5	PCT-US92-00730-1	Sequence 1, Appl 1	908	14	1.2	7320	5	PCT-US91-07141-2	Sequence 2, Appl 1
836	14	1.2	5427	5	PCT-US92-00862-1	Sequence 1, Appl 1	909	14	1.2	7362	1	US-08-455-543A-7	Sequence 7, Appl 1
837	14	1.2	5433	3	US-08-929-329-1	Sequence 1, Appl 1	910	14	1.2	7362	1	US-08-193-078B-7	Sequence 7, Appl 1
838	14	1.2	5455	1	US-08-342-930-1	Sequence 1, Appl 1	911	14	1.2	7362	2	US-08-223-305C-7	Sequence 7, Appl 1
839	14	1.2	5467	2	US-07-745-206A-12	Sequence 12, Appl 1	912	14	1.2	7362	2	US-08-149-097D-7	Sequence 7, Appl 1
840	14	1.2	5467	2	US-07-695-564-4	Sequence 4, Appl 1	913	14	1.2	7362	3	US-08-949-386-7	Sequence 7, Appl 1
841	14	1.2	5499	1	US-08-241-387-2	Sequence 2, Appl 1	914	14	1.2	7362	3	US-08-450-565-7	Sequence 7, Appl 1
842	14	1.2	5502	3	US-08-836-134-1	Sequence 1, Appl 1	915	14	1.2	7362	4	US-08-984-708A-7	Sequence 7, Appl 1
843	14	1.2	5502	3	US-07-695-564-2	Sequence 2, Appl 1	916	14	1.2	7364	4	US-09-268-165-5	Sequence 5, Appl 1
844	14	1.2	5629	1	US-08-241-387-2	Sequence 2, Appl 1	917	14	1.2	7394	4	US-09-268-165-5	Sequence 5, Appl 1
845	14	1.2	5629	1	US-09-029-213B-7	Sequence 7, Appl 1	918	14	1.2	7394	4	US-08-440-787A-6	Sequence 6, Appl 1
846	14	1.2	5820	4	US-07-745-206A-6	Sequence 6, Appl 1	919	14	1.2	7394	5	PCT-US91-07141-6	Sequence 6, Appl 1
847	14	1.2	5904	1	US-08-455-543A-3	Sequence 3, Appl 1	920	14	1.2	7394	5	US-08-367-685-6	Sequence 6, Appl 1
848	14	1.2	5904	2	US-08-193-078B-3	Sequence 3, Appl 1	921	14	1.2	7409	1	US-08-440-787A-4	Sequence 6, Appl 1
849	14	1.2	5904	2	US-08-223-305C-3	Sequence 3, Appl 1	922	14	1.2	7409	1	US-08-367-685-4	Sequence 6, Appl 1
850	14	1.2	5904	2	US-08-149-097D-3	Sequence 3, Appl 1	923	14	1.2	7409	5	PCT-US91-07141-4	Sequence 4, Appl 1
851	14	1.2	5904	2	US-08-311-363-6	Sequence 6, Appl 1	924	14	1.2	7445	1	US-08-464-136-1	Sequence 1, Appl 1
852	14	1.2	6152	1	US-08-557-139-1	Sequence 1, Appl 1	925	14	1.2	7445	1	US-08-440-787A-3	Sequence 1, Appl 1
853	14	1.2	6203	4	US-09-134-218-3	Sequence 3, Appl 1	926	14	1.2	7445	2	US-08-349-131-1	Sequence 2, Appl 1
854	14	1.2	6232	4	US-08-456-200B-11	Sequence 11, Appl 1	927	14	1.2	7445	3	US-08-470-297A-1	Sequence 3, Appl 1
855	14	1.2	6291	3	US-08-817-188-4	Sequence 4, Appl 1	928	14	1.2	7445	3	US-08-367-685-3	Sequence 3, Appl 1
856	14	1.2	6376	1	US-08-104-073-5	Sequence 5, Appl 1	929	14	1.2	7445	5	PCT-US91-07141-3	Sequence 3, Appl 1
857	14	1.2	6529	4	US-08-789-329C-1	Sequence 1, Appl 1	930	14	1.2	7502	3	US-08-969-644-6	Sequence 6, Appl 1
858	14	1.2	6529	4	US-08-949-386-3	Sequence 3, Appl 1	931	14	1.2	7502	3	US-08-444-189-6	Sequence 6, Appl 1
859	14	1.2	6575	3	US-08-450-562-3	Sequence 3, Appl 1	932	14	1.2	7502	3	US-08-464-136-1	Sequence 6, Appl 1
860	14	1.2	6575	3	US-08-984-709A-3	Sequence 3, Appl 1	933	14	1.2	7502	3	US-08-444-189-6	Sequence 6, Appl 1
861	14	1.2	6575	4	US-08-727-034-1	Sequence 1, Appl 1	934	14	1.2	7502	4	US-08-464-136-1	Sequence 6, Appl 1
862	14	1.2	6639	1	US-08-727-034-1	Sequence 1, Appl 1	935	14	1.2	7557	2	US-08-349-131-4	Sequence 4, Appl 1
863	14	1.2	6642	2	US-08-380-403A-4	Sequence 4, Appl 1	936	14	1.2	7557	5	PCT-US91-07149-4	Sequence 4, Appl 1
864	14	1.2	6642	2	US-08-895-628-4	Sequence 4, Appl 1	937	14	1.2	7557	5	US-08-287-959-2	Sequence 5, Appl 1
865	14	1.2	6693	2	US-08-147-777-2	Sequence 2, Appl 1	938	14	1.2	7573	1	US-08-867-941-6	Sequence 6, Appl 1
866	14	1.2	6693	2	US-08-452-872-2	Sequence 2, Appl 1	939	14	1.2	7635	1	US-08-455-543A-1	Sequence 2, Appl 1
867	14	1.2	6693	3	PCT-US93-03965-2	Sequence 2, Appl 1	940	14	1.2	7635	1	US-08-455-543A-23	Sequence 23, Appl 1
868	14	1.2	6725	3	US-08-949-386-36	Sequence 36, Appl 1	941	14	1.2	7635	2	US-08-193-078B-1	Sequence 1, Appl 1
869	14	1.2	6725	3	US-08-450-562-36	Sequence 36, Appl 1	942	14	1.2	7635	2	US-08-193-078B-29	Sequence 29, Appl 1
870	14	1.2	6725	3	US-08-984-709A-36	Sequence 36, Appl 1	943	14	1.2	7635	2	US-08-223-305C-1	Sequence 1, Appl 1
871	14	1.2	6725	3	US-08-984-709A-36	Sequence 36, Appl 1	944	14	1.2	7635	2	US-08-223-305C-23	Sequence 23, Appl 1
872	14	1.2	6743	3	US-08-932-280-1	Sequence 1, Appl 1	945	14	1.2	7635	2	US-08-149-097D-1	Sequence 1, Appl 1
873	14	1.2	6843	1	US-08-727-034-6	Sequence 6, Appl 1	946	14	1.2	7635	3	US-08-949-386-1	Sequence 3, Appl 1
874	14	1.2	6854	1	US-08-468-036-4	Sequence 4, Appl 1	947	14	1.2	7635	3	US-08-450-562-1	Sequence 1, Appl 1
875	14	1.2	6854	2	US-08-376-080-1	Sequence 1, Appl 1	948	14	1.2	7635	4	US-08-867-941-1	Sequence 1, Appl 1
876	14	1.2	6946	4	US-09-316-080-1	Sequence 1, Appl 1	949	14	1.2	7641	2	US-08-867-941-6	Sequence 6, Appl 1
877	14	1.2	6961	1	US-08-727-034-2	Sequence 2, Appl 1	950	14	1.2	7641	2	US-08-867-941-6	Sequence 6, Appl 1
878	14	1.2	6969	1	US-08-276-594A-1	Sequence 1, Appl 1	951	14	1.2	7641	4	US-09-074-658-6	Sequence 6, Appl 1
879	14	1.2	7011	4	US-09-268-163-9	Sequence 9, Appl 1	952	14	1.2	7641	4	US-09-074-658-6	Sequence 6, Appl 1
880	14	1.2	7032	4	US-09-324-867-1	Sequence 1, Appl 1	953	14	1.2	7650	2	US-08-867-941-1	Sequence 1, Appl 1
881	14	1.2	7056	1	US-08-121-202-1	Sequence 1, Appl 1	954	14	1.2	7650	2	US-08-867-941-1	Sequence 1, Appl 1
882	14	1.2	7083	4	US-09-198-839-1	Sequence 1, Appl 1	955	14	1.2	7650	4	US-09-074-658-1	Sequence 1, Appl 1
883	14	1.2	7125	1	US-07-745-206A-1	Sequence 1, Appl 1	956	14	1.2	7650	4	US-09-074-658-1	Sequence 1, Appl 1
884	14	1.2	7125	2	US-08-311-363-1	Sequence 1, Appl 1	957	14	1.2	7652	1	US-07-590-988A-1	Sequence 1, Appl 1
885	14	1.2	7143	4	US-09-381-862-4	Sequence 4, Appl 1	958	14	1.2	7680	4	US-09-210-748A-3	Sequence 3, Appl 1
886	14	1.2	7173	2	US-08-453-543A-8	Sequence 8, Appl 1	959	14	1.2	7729	2	US-08-464-136-3	Sequence 3, Appl 1
887	14	1.2	7175	2	US-08-193-078B-8	Sequence 8, Appl 1	960	14	1.2	7729	2	US-08-349-131-3	Sequence 3, Appl 1
888	14	1.2	7175	2	US-08-223-305C-8	Sequence 8, Appl 1	961	14	1.2	7729	2	US-08-470-297A-3	Sequence 3, Appl 1
889	14	1.2	7175	2	US-08-149-097D-8	Sequence 8, Appl 1	962	14	1.2	7729	5	PCT-US91-07149-3	Sequence 5, Appl 1
890	14	1.2	7175	3	US-08-949-386-8	Sequence 8, Appl 1	963	14	1.2	8056	1	US-09-136-605-14	Sequence 14, Appl 1
891	14	1.2	7175	3	US-08-450-562-8	Sequence 8, Appl 1	964	14	1.2	8056	1	US-08-306-691B-41	Sequence 41, Appl 1
892	14	1.2	7175	4	US-08-984-709A-8	Sequence 8, Appl 1	965	14	1.2	8082	1	US-08-187-785-1	Sequence 1, Appl 1
893	14	1.2	7177	4	US-09-268-163-7	Sequence 7, Appl 1	966	14	1.2	8082	5	PCT-US93-06251-28	Sequence 28, Appl 1
894	14	1.2	7266	3	US-08-713-118-1	Sequence 1, Appl 1	967	14	1.2	8118	1	US-08-464-136-5	Sequence 5, Appl 1
895	14	1.2	7266	3	US-09-453-007-1	Sequence 1, Appl 1	968	14	1.2	8118	2	US-08-349-131-5	Sequence 5, Appl 1
896	14	1.2	7294	1	US-08-440-787A-1	Sequence 1, Appl 1	969	14	1.2	8118	3	US-08-470-297A-5	Sequence 5, Appl 1
897	14	1.2	7294	1	US-08-440-787A-5	Sequence 5, Appl 1	970	14	1.2	8241	6	PCT-US91-07149-5	Sequence 6, Appl 1
898	14	1.2	7294	4	US-08-367-685-1	Sequence 1, Appl 1	971	14	1.2	8241	6	US-08-462-014-2	Sequence 2, Appl 1
899	14	1.2	7294	4	US-08-367-685-5	Sequence 5, Appl 1	972	14	1.2	8299	3	US-08-923-137-3	Sequence 3, Appl 1
900	14	1.2	7294	5	PCT-US91-07141-1	Sequence 1, Appl 1	973	14	1.2	8299	4	US-08-923-137-3	Sequence 3, Appl 1
901	14	1.2	7294	5	PCT-US91-07141-5	Sequence 5, Appl 1	974	14	1.2	8299	4	US-08-923-137-3	Sequence 3, Appl 1
902	14	1.2	7317	1	US-08-464-136-2	Sequence 2, Appl 1	975	14	1.2	8299	4	US-08-923-137-3	Sequence 3, Appl 1
903	14	1.2	7317	2	US-08-349-131-2	Sequence 2, Appl 1	976	14	1.2	8321	3	US-08-680-506-1	Sequence 1, Appl 1

```
977 14 1.2 8341 2 US-08-737-825-1 Sequence 1, Appli
978 14 1.2 8509 1 US-08-462-014-1 Sequence 1, Appli
979 14 1.2 8509 4 US-08-973-334-4 Sequence 4, Appli
980 14 1.2 8509 4 US-09-563-869A-4 Sequence 4, Appli
981 14 1.2 8640 1 US-08-570-311-28 Sequence 28, Appli
982 14 1.2 8967 1 US-08-366-851A-1 Sequence 1, Appli
983 14 1.2 9009 1 US-07-864-004B-3 Sequence 3, Appli
984 14 1.2 9009 1 US-08-251-937A-3 Sequence 3, Appli
985 14 1.2 9009 1 US-08-212-133A-1 Sequence 1, Appli
986 14 1.2 9009 1 US-08-474-503-1 Sequence 1, Appli
987 14 1.2 9009 2 US-08-670-707A-1 Sequence 1, Appli
988 14 1.2 9009 4 US-09-037-601-1 Sequence 1, Appli
989 14 1.2 9009 5 PCT-US93-03275-3 Sequence 3, Appli
990 14 1.2 9009 5 PCT-US94-13200-1 Sequence 1, Appli
991 14 1.2 9046 1 US-08-227-536-1 Sequence 1, Appli
992 14 1.2 9046 5 PCT-US95-04682-1 Sequence 1, Appli
993 14 1.2 9354 1 US-08-683-839B-2 Sequence 2, Appli
994 14 1.2 9370 1 US-08-320-559-27 Sequence 27, Appli
995 14 1.2 9370 3 US-08-545-860D-27 Sequence 27, Appli
996 14 1.2 9370 5 PCT-US94-04496-27 Sequence 27, Appli
997 14 1.2 9391 1 US-08-320-559-25 Sequence 25, Appli
998 14 1.2 9391 3 US-08-545-860D-25 Sequence 25, Appli
999 14 1.2 9391 5 PCT-US94-04496-25 Sequence 25, Appli
1000 14 1.2 9723 1 US-08-083-590A-21 Sequence 21, Appli
```

## ALIGNMENTS

```
RESULT 1
US-08-451-822A-15/c
; Sequence 15, Application US/08451822A
; Patent No. 5863888
```

## GENERAL INFORMATION:

```
APPLICANT: Dionne, Craig A
APPLICANT: Crumley, Greg
APPLICANT: Jaye, Michael C
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: Fibroblast Growth Factor Receptors
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Legal Department
STREET: 500 Arcola Road
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
```

## COMPUTER READABLE FORM:

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451, 822A
FILING DATE: 26-MAY-1995
```

## CLASSIFICATION: 435

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,430
FILING DATE: 14-OCT-1994
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/934,372
FILING DATE: 21-AUG-1992
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
```

## ATTORNEY/AGENT INFORMATION:

```
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A0496E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 15:
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-451-822A-15
```

```
Query Match 1.6%; Score 18; DB 2; Length 3416;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 38 gtctgctgtgtcgtcgc 55
Db 3347 GTCGCTGTGCTGCTGC 3330
```

## RESULT 2

```
US-08-323-430-15/c
; Sequence 15, Application US/08323430
; Patent No. 6344546
```

## GENERAL INFORMATION:

```
APPLICANT: Dionne, Craig A
APPLICANT: Crumley, Greg
APPLICANT: Jaye, Michael C
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: Fibroblast Growth Factor Receptors
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Legal Department
STREET: 500 Arcola Road
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
```

## COMPUTER READABLE FORM:

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,430
FILING DATE:
CLASSIFICATION: 435
```

## PRIOR APPLICATION DATA:

```
APPLICATION NUMBER: US to be assigned
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
```

## ATTORNEY/AGENT INFORMATION:

```
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A0496E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 454-3817
TELEFAX: (215) 454-3808
INFORMATION FOR SEQ ID NO: 15:
```

## SEQUENCE CHARACTERISTICS:

```
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-323-430-15
```

```
Query Match 1.6%; Score 18; DB 4; Length 3416;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 38 gtctgctgtgtcgtcgc 55
|||||
```



Db 3347 GTCGTGCTGCTGCTGC 3330

## RESULT 3

US-09-150-460B-3  
; Sequence 3, Application US/09150460B  
; Patent No. 6190882  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Cheng-Chi  
; APPLICANT: Albrecht, Urs  
; APPLICANT: Eichele, Gregor  
; APPLICANT: Sun, Zhong Sheng  
; TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene  
; FILE REFERENCE: D6039  
; CURRENT APPLICATION NUMBER: US/09/150,460B  
; CURRENT FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/058,256  
; PRIOR FILING DATE: 1997-09-09  
; NUMBER OF SEQ ID NOS: 21  
; SEQ ID NO 3  
; LENGTH: 6614  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of 6.6 kb cDNA of human RIGUI  
US-09-150-460B-3

Query Match 1.6%; Score 18; DB 4; Length 6614;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ctgtcgtcgtcgtctgtct 60  
|||||  
Db 4194 ctgtcgtcgtcgtctgtct 4211

## RESULT 4

US-08-932-978-1  
; Sequence 1, Application US/08932978  
; Patent No. 5885804  
; GENERAL INFORMATION:  
; APPLICANT: Zalacain, Magdalena  
; APPLICANT: Brown, James R.  
; TITLE OF INVENTION: NOVEL PHO  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/932,978  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dickinson, Todd O  
; REGISTRATION NUMBER: 28,354  
; REFERENCE/DOCKET NUMBER: GM0100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2252  
; TELEFAX: 215-994-2222  
; TELEX:

; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 969 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-932-978-1

Query Match 1.5%; Score 17; DB 2; Length 969;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 gcgtgaattatcaaaa 136  
|||||  
Db 630 GCGTGAATTATCAAAA 646

## RESULT 5

US-08-714-918-65/c  
; Sequence 65, Application US/08714918  
; Patent No. 6037123  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,918  
; FILING DATE: September 13, 1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 222/005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2976 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-714-918-65

Query Match 1.5%; Score 17; DB 3; Length 2976;

Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 797 tgatgaacacctgaaa 813  
|||||  
DB 2333 TGATGAACAACCTGAAA 2317

RESULT 6  
US-09-265-315-65/c  
; Sequence 65, Application US/09265315  
; Patent No. 6187541  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ying J.  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,315  
; FILING DATE: March 9, 1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/714,918  
; FILING DATE: September 13, 1996  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Waiburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 240/247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEFAX: 67-3510  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2976 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-265-315-65

Query Match 1.5%; Score 17; DB 4; Length 2976;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 797 tgatgaacacctgaaa 813  
|||||  
DB 2333 TGATGAACAACCTGAAA 2317

RESULT 7  
US-09-265-315-65/c  
; Sequence 65, Application US/09265315  
; Patent No. 6187541  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ying J.  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,315  
; FILING DATE: March 9, 1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/714,918  
; FILING DATE: September 13, 1996  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Waiburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 240/247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEFAX: 67-3510  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2976 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-265-315-65

Query Match 1.5%; Score 17; DB 4; Length 2976;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 797 tgatgaacacctgaaa 813  
|||||  
DB 2333 TGATGAACAACCTGAAA 2317

RESULT 8  
US-09-266-417-65/c  
; Sequence 65, Application US/09266417  
; Patent No. 628588  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret

APPLICANT: Lee, Ving J.  
APPLICANT: Malouin, Francois  
APPLICANT: Martin, Patrick K.  
APPLICANT: Schmid, Molly B.  
APPLICANT: Sun, Dongxu  
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
TITLE OF INVENTION: TARGET GENES  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/266,417  
FILING DATE: March 9, 1999  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/714, 918  
FILING DATE: September 13, 1996  
APPLICATION NUMBER: 60/009,102  
FILING DATE: December 22, 1995  
APPLICATION NUMBER: 60/003,798  
FILING DATE: September 15, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 240/248  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2976 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-266-417-65

Query Match 1.5%; Score 17; DB 4; Length 2976;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 tgatgaacaactga 813  
|||||  
Db 2333 TGATGACACACTGAAA 2317

RESULT 9  
US-08-070-301-2  
Sequence 2, Application US/08070301  
Patent No. 5871995  
GENERAL INFORMATION:  
APPLICANT: IIDA, Toshio  
APPLICANT: KAMINUMA, Toshihiko  
APPLICANT: FUSE, Yuka  
APPLICANT: TAJIMA, Masahiko  
APPLICANT: YAMAGI, Mitsuo  
APPLICANT: OKAMOTO, Hiroshi  
APPLICANT: KISHIMOTO, Jiro  
APPLICANT: IFUKU, Ohji

APPLICANT: KATO, Ichiro  
TITLE OF INVENTION: ENZYME PARTICIPATING IN C-TERMINAL  
TITLE OF INVENTION: AMIDATION, AND METHOD OF PREPARING SAME AND USE THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wegner, Cantor, Mueller & Player, P.C.  
STREET: 1233 20th Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036-8218  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/070,301  
FILING DATE: 24-MAY-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-209687  
FILING DATE: 15-AUG-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-181933  
FILING DATE: 31-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-76331  
FILING DATE: 26-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-106412  
FILING DATE: 24-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-205475  
FILING DATE: 02-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: P-450-22830  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-040  
TELEFAX: (202) 835-0605  
TELEX: 440706  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Horse  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 11..3070  
US-08-070-301-2

Query Match 1.5%; Score 17; DB 2; Length 6638;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 gatggcctgagatata 334  
|||||  
Db 1514 GATTGGCTGGAGATATA 1530

RESULT 10  
US-08-846-111D-15/c  
Sequence 15, Application US/08846111D  
Patent No. 6017705  
GENERAL INFORMATION:  
APPLICANT: Lurquin, Christophe; Brasseur, Francis;

APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Are  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect For DOS 6.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/846.111D  
FILING DATE: 25-APRIL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/658,578  
FILING DATE: 5-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/403,388  
FILING DATE: 14-MARCH-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6017705man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5444.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40352 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-846-111D-15

Query Match 1.5%; Score 17; DB 3; Length 40352;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 200 acaactgaagacctg 216  
|||||  
Db 13873 ACAACCTGAAGACCTG 13857

RESULT 11  
US-08-920-422-17/c  
Sequence 17, Application US/08920422A  
Patent No. 6255473  
GENERAL INFORMATION:  
APPLICANT: Vitek, Michael P.  
APPLICANT: Mitsuda, No. 6255473iak1  
APPLICANT: Roses, Allen D.  
TITLE OF INVENTION: Presentin-1 Gene Promoter  
FILE REFERENCE: VITEKPRESENTIN  
CURRENT APPLICATION NUMBER: US/08/920.422A  
FILING DATE: 1997-08-29  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 17  
LENGTH: 48974  
TYPE: DNA  
ORGANISM: Mus musculus  
US-08-920-422-17

Query Match 1.5%; Score 17; DB 4; Length 48974;  
Best Local Similarity 100.0%; Pred. No. 31;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1016 taaaactgaagaa 1032  
|||||  
Db 34191 TAAACTGAAACAA 34175

RESULT 12  
US-09-651-656-2/c  
Sequence 2, Application US/09651656  
Patent No. 6340566  
GENERAL INFORMATION:  
APPLICANT: MCCUTHEN-MALONEY, SANDRA  
TITLE OF INVENTION: LAWRENCE LIVERMORE NATIONAL LABORATORY  
TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE  
TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,  
FILE REFERENCE: IL-10689  
CURRENT APPLICATION NUMBER: US/09/651.656  
CURRENT FILING DATE: 2000-08-29  
PRIOR APPLICATION NUMBER: 60/192,764  
PRIOR FILING DATE: 2000-03-28  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 723  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-651-656-2

Query Match 1.4%; Score 16; DB 4; Length 723;  
Best Local Similarity 100.0%; Pred. No. 16+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 281 aatcatcgtagtag 296  
|||||  
Db 371 AATCATCGTAGTAG 356

RESULT 13  
US-08-982-956-2  
Sequence 2, Application US/08982956  
Patent No. 5861312  
GENERAL INFORMATION:  
APPLICANT: Varshavsky, Alexander  
APPLICANT: Kwon, Yong Tae  
TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN DBRI  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Kevin M. Farrell, P.C.  
STREET: P.O. Box 999  
CITY: York Harbor  
STATE: ME  
COUNTRY: US  
ZIP: 03911  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/982.956  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farrell, Kevin M.  
REGISTRATION NUMBER: 35,505  
REFERENCE/DOCKET NUMBER: CIT-2001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (207) 363-0558  
TELEFAX: (207) 363-0528  
INFORMATION FOR SEQ ID NO: 2:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 1001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..999
US-08-982-956-2

Query Match 1.4%; Score 16; DB 2; Length 1001;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 548 atgataaagaagaac 563
Db 932 ATGATTAAGAAAAAGC 947

RESULT 14
US-09-228-317-2
Sequence 2, Application US/09228317
Patent No. 6159732
GENERAL INFORMATION:
APPLICANT: Vaishavsky, Alexander
APPLICANT: Kwon, Yong Tae
TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBRL
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: ME
COUNTRY: US
ZIP: 03911
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/228,317
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M.
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: CIT-2001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (207) 363-0558
TELEFAX: (207) 363-0528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..999
US-09-228-317-2

Query Match 1.4%; Score 16; DB 3; Length 1001;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 548 atgataaagaagaac 563
Db 932 ATGATTAAGAAAAAGC 947
```

```
RESULT 15
US-08-176-427B-9
Sequence 9, Application US/08176427B
Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablo, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1256 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1257
US-08-176-427B-9
```

```
Query Match 1.4%; Score 16; DB 1; Length 1256;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 tgaaggaactgaacc 221
Db 644 TGAAGGACCTGAACCC 659
```

Search completed: May 7, 2002, 18:11:29  
Job time: 8135 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2002, 15:09:30 ; Search time 1802.55 Seconds  
(without alignments)  
8513.516 Million cell updates/sec

Title: US-09-579-383-1  
Perfect score: 1137  
Sequence: 1 atgaatttaccgtaaataa.....tcgttgaagcctgcactaa 1137

Scoring table: OLIGO MNC  
Gapex 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estz:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_liv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	1.9	348	9	A1176234
C 2	22	1.9	348	9	A1176234 EST219815
C 3	21	1.8	585	12	A0187817
C 4	21	1.8	627	12	A0891144
C 5	20	1.8	282	9	BS151310
C 6	20	1.8	339	9	AA269349
C 7	20	1.8	370	9	A1834906
C 8	20	1.8	373	9	AA590068
C 9	20	1.8	376	10	W30259
C 10	20	1.8	383	9	AM822056
C 11	20	1.8	406	10	W69001
C 12	20	1.8	409	9	AA238758
C 13	20	1.8	413	10	BF783504
C 14	20	1.8	426	9	AA245797
C 15	20	1.8	428	9	A1042890
C 16	20	1.8	432	9	A1173373
C 17	20	1.8	432	9	AL118209

C 18	20	1.8	446	9	A1848779
C 19	20	1.8	476	9	AA743859
C 20	20	1.8	488	9	AA237577
C 21	20	1.8	498	9	AA743402
C 22	20	1.8	510	10	B1320819
C 23	20	1.8	512	12	TA196A01P
C 24	20	1.8	531	9	AA881704
C 25	20	1.8	545	12	AQ358119
C 26	20	1.8	549	12	AQ653019
C 27	20	1.8	553	12	AQ943352
C 28	20	1.8	556	12	TA323B01P
C 29	20	1.8	608	9	AA141807
C 30	20	1.8	624	12	AQ653083
C 31	20	1.8	645	9	BB662515
C 32	20	1.8	676	9	BE062386
C 33	20	1.8	678	10	B1407878
C 34	20	1.8	681	9	BB521972
C 35	20	1.8	690	9	BB653705
C 36	20	1.8	710	10	B1415854
C 37	20	1.8	711	10	BF160523
C 38	20	1.8	714	10	BG969104
C 39	20	1.8	724	9	BB667504
C 40	20	1.8	747	10	B1905631
C 41	20	1.8	759	10	BP384316
C 42	20	1.8	764	10	BG964617
C 43	20	1.8	782	10	B1157508
C 44	20	1.8	784	10	B1327637
C 45	20	1.8	794	10	B1407480
C 46	20	1.8	799	10	B1102992
C 47	20	1.8	831	10	B1218401
C 48	20	1.8	873	10	B1410511
C 49	20	1.8	888	11	AK002424
C 50	20	1.8	1091	12	CNS030EF
C 51	20	1.8	1104	10	BF787117
C 52	20	1.8	1295	10	B114965
C 53	20	1.7	213	10	BE830429
C 54	19	1.7	234	9	AV239268
C 55	19	1.7	319	9	BB143865
C 56	19	1.7	355	10	H92891
C 57	19	1.7	359	10	H92777
C 58	19	1.7	364	12	AQ770155
C 59	19	1.7	381	9	AM104487
C 60	19	1.7	384	12	AQ933957
C 61	19	1.7	387	9	AA219187
C 62	19	1.7	389	9	AV536272
C 63	19	1.7	392	12	AZ490233
C 64	19	1.7	422	9	AM124623
C 65	19	1.7	436	10	B1201479
C 66	19	1.7	437	12	CNS00P08
C 67	19	1.7	454	9	AJ284734
C 68	19	1.7	456	9	AU240425
C 69	19	1.7	459	9	AM674571
C 70	19	1.7	481	9	AV540025
C 71	19	1.7	483	10	B1077260
C 72	19	1.7	512	12	AZ173339
C 73	19	1.7	518	9	A1911783
C 74	19	1.7	528	12	AZ856192
C 75	19	1.7	559	12	AQ925396
C 76	19	1.7	568	9	AA663483
C 77	19	1.7	578	10	B1612817
C 78	19	1.7	589	9	A1812873
C 79	19	1.7	609	9	AM303371
C 80	19	1.7	616	12	AG085862
C 81	19	1.7	622	9	AV405606
C 82	19	1.7	639	9	A1594699
C 83	19	1.7	651	9	BB441046
C 84	19	1.7	654	10	B1627704
C 85	19	1.7	654	12	AG129762
C 86	19	1.7	671	9	AM646020
C 87	19	1.7	699	9	BE038084
C 88	19	1.7	702	9	A1815332
C 89	19	1.7	702	12	AG033923
C 90	19	1.7	713	12	AZ347710

A1848779	UI-M-A11-
AA743859	up66a05.x
AA237577	mx11C02.r
AA743402	up67b01.y
B1320819	sae49e10.
AA75894	T. brucei
AA881704	vx26c09.r
AQ358119	HS_5027_A
AQ653019	Sheared D
AQ943352	Sheared D
AA1493417	T. brucei
AA141807	CK02546.5
AQ653083	Sheared D
BB662515	BB662515
BE062386	RC4-BT025
B1407878	602919293
BB521972	BB521972
BB653705	BB653705
B1415854	602988050
BF160523	601766894
BG969104	602835078
BB667504	BB667504
B1905631	603167634
BP384316	602046664
BG964617	602831803
B1157508	602923530
B1327637	602979130
B1407480	602919293
B1102992	602888611
B1218401	602937785
B1410511	602965018
AK002424	Mus muscu
AL255892	Tetradodon
BF787117	602106654
B114965	602862779
BE830429	RC6-ET008
AV239268	AV239268
BB143865	BB143865
H92891	YF94C02.s1
H92777	YF94C02.r1
AQ770155	HS_2002_A
AM104487	xd87a04.x
AQ933957	RPC1-23-2
AA219187	zq16d08.r
AV536272	AV536272
AZ490233	1M0323N02
AM124623	UI-M-BH2.
B1201479	p3h04fs.r
A1084758	Arabidops
AJ284734	4A3B-AAH-
AU240425	AU240425
AM674571	bb34a04.y
AV540025	AV540025
B1077260	NXPV_094-
A1173339	SP_0123_B
A1911783	wd32g11.x
AZ856192	2M0160H02
AQ925396	RPC1-23-2
AA663483	cs51f07.s
B1612817	RM41794.5
A1812873	22A12_Pln
AM303371	xv18c07.x
AG085862	Par. trogl
AV405606	AV405606
A1594699	AL594699
BB441046	BB441046
B1627704	RH69082.5
AG129762	Par. trogl
AM646020	cm60a06.w
BE038084	AA08F08.A
A1815332	SMOvAPCHI
AG033923	Par. trogl
AZ347710	IM0083J19

91	19	1.7	743	10	BF233940	BF233940 602024421	C 164	18	1.6	590	10	BI345036	BI345036 373732 MA
C 92	19	1.7	744	12	AQ288319	AQ288319 nxb0032L	C 165	18	1.6	598	10	BI682434	BI682434 463516 MA
C 93	19	1.7	769	12	CNS04EMP	AL288579 Tetradon	C 166	18	1.6	600	10	BG807186	BG807186 2052-26 M
C 94	19	1.7	832	12	AZ535507	AZ535507 ENTCC68TF	C 167	18	1.6	610	10	BI357135	BI357135 RE43523.5
C 95	19	1.7	873	12	CNS02PMZ	AL217628 Tetradon	C 168	18	1.6	610	10	BE712404	BE712404 QV2-HT069
C 96	19	1.7	931	12	CNS02PCA	AL207811 Tetradon	C 169	18	1.6	611	10	BM262058	BM262058 daq40b09.
C 97	19	1.7	1003	12	CNS024Z7	AL181420 Tetradon	C 170	18	1.6	612	9	AW949428	AW949428 EST361498
C 98	18	1.6	207	10	BI006414	BI006414 MRL-RT003	C 171	18	1.6	617	9	AW949426	AW949426 EST361496
C 99	18	1.6	253	10	BG191540	BG191540 RST10634	C 172	18	1.6	621	12	AZ818350	AZ818350 2M0088N18
C 100	18	1.6	265	10	AWA35478	AWA35478 UT-R-BOP	C 173	18	1.6	622	10	BI357586	BI357586 RE44110.5
C 101	18	1.6	273	9	BB544410	BB544410 BB544410	C 174	18	1.6	622	10	BI362003	BI362003 RE46452.15
C 102	18	1.6	281	10	BE500504	BE500504 WHE0987-0	C 175	18	1.6	625	9	BE660052	BE660052 856 Gmax
C 103	18	1.6	315	9	AJ280504	AJ280504 4A3A-AAO-	C 176	18	1.6	629	9	AV651738	AV651738 AV651738
C 104	18	1.6	337	12	AQ062002	AQ062002 CIT-HSP-2	C 177	18	1.6	631	10	BE660051	BE660051 1-C6 Gmax
C 105	18	1.6	364	10	BI180405	BI180405 949031E02	C 178	18	1.6	636	10	BI622572	BI622572 RH54466.5
C 106	18	1.6	380	10	BI622807	BI622807 RH54730.5	C 179	18	1.6	636	10	BE712434	BE712434 QV2-HT069
C 107	18	1.6	382	9	AA387765	AA387765 V558B06.r	C 180	18	1.6	638	10	BE712405	BE712405 QV2-HT069
C 108	18	1.6	390	9	AA490022	AA490022 ab02b04.s	C 181	18	1.6	638	12	AZ343817	AZ343817 IM0077E03
C 109	18	1.6	392	12	AQ486017	AQ486017 RPCI-11-2	C 182	18	1.6	642	9	AW305549	AW305549 f160c12.y
C 110	18	1.6	395	9	BB800724	BB800724 BP800724	C 183	18	1.6	643	10	BG086723	BG086723 LERAD35TF
C 111	18	1.6	400	9	AWA61441	AWA61441 BP230005A	C 184	18	1.6	646	12	AG095532	AG095532 LERAD35TF
C 112	18	1.6	402	10	R70205	R70205 v180b12.r1	C 185	18	1.6	646	12	AZ955245	AZ955245 2M0221G13
C 113	18	1.6	403	9	AA834917	AA834917 aj37e12.s	C 186	18	1.6	651	12	AG133350	AG133350 Pan LT091
C 114	18	1.6	409	10	BE620895	BE620895 601483747	C 187	18	1.6	653	10	BI163572	BI163572 RE03006.5
C 115	18	1.6	412	9	AW159225	AW159225 za62e02.x	C 188	18	1.6	664	9	AV681532	AV681532 AV681532
C 116	18	1.6	415	9	AW429933	AW429933 68542 MAR	C 189	18	1.6	673	10	BG706936	BG706936 602672123
C 117	18	1.6	417	9	AW830988	AW830988 sm21a03.Y	C 190	18	1.6	712	9	AV726725	AV726725 AV726725
C 118	18	1.6	421	12	AZ487275	AZ487275 IM0316117	C 191	18	1.6	717	12	BH244864	BH244864 AUTJA78TF
C 119	18	1.6	437	12	BH255971	BH255971 RM0109AM0	C 192	18	1.6	729	9	AI906341	AI906341 IL-BT108-
C 120	18	1.6	438	12	AZ589047	AZ589047 IM0397G18	C 193	18	1.6	734	12	CNS05CRJ	AL331480 Tetradon
C 121	18	1.6	440	10	BE863083	BE863083 UT-M-BHO-	C 194	18	1.6	736	10	BF786714	BF786714 602111477
C 122	18	1.6	442	9	AA235910	AA235910 z543c12.r	C 195	18	1.6	766	12	AZ416858	AZ416858 IM0192116
C 123	18	1.6	446	12	AZ295057	AZ295057 RPCI-23-1	C 196	18	1.6	786	10	BI092661	BI092661 602858449
C 124	18	1.6	448	9	AI627656	AI627656 t81c08.x	C 197	18	1.6	789	12	AG093746	AG093746 Pan tlog1
C 125	18	1.6	448	9	AV440616	AV440616 AV440616	C 198	18	1.6	810	12	AQ060685	AQ060685 nbebd015F
C 126	18	1.6	450	12	AZ331087	AZ331087 IM0056D16	C 199	18	1.6	812	12	BH092860	BH092860 RPCI-24-3
C 127	18	1.6	453	12	AQ0802835	AQ0802835 HS_3186-A	C 200	18	1.6	817	9	AF074799	AF074799 AF074799
C 128	18	1.6	454	12	AQ060044	AQ060044 CIT-HSP-2	C 201	18	1.6	817	10	BI660191	BI660191 603302567
C 129	18	1.6	455	12	AQ481220	AQ481220 RPCI-11-2	C 202	18	1.6	822	12	AQ955331	AQ955331 LERAD35TF
C 130	18	1.6	477	12	B43843	B43843 HS-1058-A2-	C 203	18	1.6	825	12	CNS04KGI	AL294795 Tetradon
C 131	18	1.6	479	9	AV596685	AV596685 AV596685	C 204	18	1.6	832	9	AF074804	AF074804 AF074804
C 132	18	1.6	482	10	BF769079	BF769079 QV3-TT000	C 205	18	1.6	834	10	BI765848	BI765848 603049874
C 133	18	1.6	487	9	AA664355	AA664355 at25h05.s	C 206	18	1.6	834	12	AQ744814	AQ744814 HS_5506-A
C 134	18	1.6	488	9	AW193590	AW193590 xm09h10.x	C 207	18	1.6	844	12	CNS0454G	AL274921 Tetradon
C 135	18	1.6	489	12	BG673531	BG673531 DRNCFH09	C 208	18	1.6	858	9	AI403531	AI403531 CH22991.5
C 136	18	1.6	502	10	BH442152	BH442152 BOHV64TF	C 209	18	1.6	872	12	BH131071	BH131071 ENTNN12TF
C 137	18	1.6	507	10	BG511299	BG511299 s8d16b02.	C 210	18	1.6	879	12	AZ693029	AZ693029 ENTLD94TF
C 138	18	1.6	515	10	BE329989	BE329989 s8d72a11.y	C 211	18	1.6	935	10	BE549386	BE549386 601075773
C 139	18	1.6	517	10	BE721156	BE721156 188212.MA	C 212	18	1.6	937	12	CNS043D8	AL272645 Tetradon
C 140	18	1.6	519	12	AQ688657	AQ688657 n0xb0078E	C 213	18	1.6	941	10	BF532233	BF532233 602073134
C 141	18	1.6	519	12	AZ288853	AZ288853 RPCI-23-1	C 214	18	1.6	997	12	CNS04AR1	AL303323 Tetradon
C 142	18	1.6	520	12	BH198866	BH198866 TC3-6F21	C 215	18	1.6	1006	10	BI086478	BI086478 602849748
C 143	18	1.6	522	10	BG553721	BG553721 da627a11.	C 216	18	1.6	1036	10	BG390165	BG390165 602415687
C 144	18	1.6	524	10	BF018702	BF018702 ux97b02.x	C 217	18	1.6	1085	10	BI410118	BI410118 602964050
C 145	18	1.6	524	12	AQ0824284	AQ0824284 HS_3249-B	C 218	18	1.6	1190	10	BG844746	BG844746 1024007F0
C 146	18	1.6	528	9	AI042703	AI042703 uc76f07.x	C 219	18	1.6	1325	10	BI331809	BI331809 602982380
C 147	18	1.6	529	10	BI608157	BI608157 RML2334.5	C 220	18	1.6	2219	11	AKO11270	AKO11270 Mus muscu
C 148	18	1.6	532	12	AQ080825	AQ080825 HS_5258.B	C 221	18	1.6	3952	11	AKO14555	AKO14555 Mus muscu
C 149	18	1.6	539	9	AJ280428	AJ280428 4A3A-AAP-	C 222	17	1.5	81	9	AA897476	AA897476 a148e11.s
C 150	18	1.6	540	12	AZ145902	AZ145902 SP_0015-B	C 223	17	1.5	87	12	AZ776811	AZ776811 2M0010113
C 151	18	1.6	542	12	AZ067610	AZ067610 RPCI-23-4	C 224	17	1.5	146	9	AV174887	AV174887 AV174887
C 152	18	1.6	543	9	BE124087	BE124087 EST394212	C 225	17	1.5	154	10	T20190	T20190 D165F Heart
C 153	18	1.6	544	9	AI041534	AI041534 ovr82b07.x	C 226	17	1.5	157	9	AW318993	AW318993 UN09607.y
C 154	18	1.6	546	12	AQ083267	AQ083267 HS_5503.B	C 227	17	1.5	174	12	AQ276846	AQ276846 CITBI-EL-
C 155	18	1.6	549	9	AA183718	AA183718 mC23h06.T	C 228	17	1.5	178	10	BM166184	BM166184 EST568707
C 156	18	1.6	552	9	AJ281632	AJ281632 4A3A-P8C6	C 229	17	1.5	186	12	B37915	B37915 HS_1046-A2-
C 157	18	1.6	555	10	BI615400	BI615400 RH45015.5	C 230	17	1.5	191	9	AV624058	AV624058 AV624058
C 158	18	1.6	555	10	BI615400	BI615400 RH45015.5	C 231	17	1.5	198	9	BB405044	BB405044 BB405044
C 159	18	1.6	562	9	BI658268	BI658268 fc20g10.Y	C 232	17	1.5	207	10	BM421440	BM421440 V014D07 O
C 160	18	1.6	563	9	BB695595	BB695595 BB695595	C 233	17	1.5	212	9	AV263336	AV263336 AV263336
C 161	18	1.6	567	12	BH554683	BH554683 BOHTE62TF	C 234	17	1.5	215	10	BG602891	BG602891 EST501981
C 162	18	1.6	575	9	AA141048	AA141048 CK01116.C	C 235	17	1.5	215	10	BB035921	BB035921 BB035921
C 163	18	1.6	575	12	AQ734966	AQ734966 HS_3085_B	C 236	17	1.5	216	9	BB214004	BB214004 BB214004



C 237	17	1.5	219	9	AV087749	AV087749	310	17	1.5	361	10	T26683	T26683 T047 MYAT4
C 238	17	1.5	219	9	BB073298	BB073298	311	17	1.5	362	9	AW227209	AW227209 umf7e07.y
C 239	17	1.5	232	10	BG626388	CC-esf1cL	312	17	1.5	366	9	AJ283431	AJ283431 4A3B-AAB-
C 240	17	1.5	233	9	AV369395	AV369395	313	17	1.5	367	9	AI810136	AI810136 wfe5h09.x
C 241	17	1.5	236	9	BB212611	BB212611	314	17	1.5	368	10	BE864213	BE864213 UI-M-BH1-
C 242	17	1.5	244	9	AV246831	AV246831	315	17	1.5	369	12	BH220676	BH220676 1006096C0
C 243	17	1.5	245	9	BB019686	BB019686	316	17	1.5	373	9	AA318102	AA318102 EST70171
C 244	17	1.5	247	10	BI433041	BI433041	317	17	1.5	374	9	AV682452	AV682452 AV682452
C 245	17	1.5	250	10	BI754419	BI754419	318	17	1.5	377	10	BM246803	BM246803 K0743B06-
C 246	17	1.5	251	9	AV563105	AV563105	319	17	1.5	377	10	BE485174	BE485174 172102 BA
C 247	17	1.5	252	9	AV355936	AV355936	320	17	1.5	378	10	BI568421	BI568421 RH39864.5
C 248	17	1.5	253	9	AV869532	AV869532	321	17	1.5	379	10	D60823	D60823 HUM13B04A
C 249	17	1.5	255	9	BB608289	BB608289	322	17	1.5	380	9	AA451151	AA451151 vF88C03.r
C 250	17	1.5	263	12	BH014439	UTG6BL77TH	323	17	1.5	380	9	BB800684	BB800684 BB800684
C 251	17	1.5	267	10	BE624836	BE624836	324	17	1.5	381	10	H96089	H96089 yv98e12..s1
C 252	17	1.5	269	9	BB083608	BB083608	325	17	1.5	383	10	BF766142	BF766142 IL2-CS007
C 253	17	1.5	270	9	AA893954	AA893954	326	17	1.5	384	9	AA889167	AA889167 ak25e06.s
C 254	17	1.5	271	9	AM086481	xc74f11.x	327	17	1.5	384	10	BF787716	BF787716 MR1-KT005
C 255	17	1.5	273	9	AT563535	EST00659	328	17	1.5	385	9	AI998537	AI998537 701546087
C 256	17	1.5	274	9	AM319589	AM319589	329	17	1.5	385	10	BG603969	BG603969 EST503059
C 257	17	1.5	275	9	BB370551	BB370551	330	17	1.5	386	9	AI956576	AI956576 ul78h12.y
C 258	17	1.5	277	10	BI565449	BI565449	331	17	1.5	386	9	BB846189	BB846189 BB846189
C 259	17	1.5	278	9	BB150575	BB150575	332	17	1.5	389	12	AQ089359	AQ089359 HS_318B_B
C 260	17	1.5	279	9	BB484459	BB484459	333	17	1.5	390	9	AA682873	AA682873 z115h10.s
C 261	17	1.5	279	10	BF454095	BF454095	334	17	1.5	390	9	BB669162	BB669162 BB669162
C 262	17	1.5	280	9	AA795031	AA795031	335	17	1.5	390	10	BG668501	BG668501 DRNAALD05
C 263	17	1.5	280	9	AV902947	AV902947	336	17	1.5	392	9	AJ280668	AJ280668 4A3A-AAS-
C 264	17	1.5	280	12	A2751102	RPCT-24-1	337	17	1.5	393	9	BE155508	BE155508 PM1-HT035
C 265	17	1.5	281	9	BB328252	BB328252	338	17	1.5	393	10	H60920	H60920 yT14d05.r1
C 266	17	1.5	285	9	AA486409	73342 MAR	339	17	1.5	394	12	AZ939550	AZ939550 2M0198N03
C 267	17	1.5	287	9	BB371042	BB371042	340	17	1.5	395	10	BM336214	BM336214 MEST191-A
C 268	17	1.5	290	9	BB080130	BB080130	341	17	1.5	396	12	AQ178354	AQ178354 HS_-2220_A
C 269	17	1.5	290	9	BB128545	BB128545	342	17	1.5	397	12	AQ136268	AQ136268 HS_-3056_B
C 270	17	1.5	294	10	BF748031	MR2-BN038	343	17	1.5	398	10	BG726422	BG726422 sad44d12.
C 271	17	1.5	294	10	BI028715	CMO-MT018	344	17	1.5	400	9	AA798061	AA798061 vW32e02.r
C 272	17	1.5	294	10	BF453154	BF453154	345	17	1.5	401	9	AV442183	AV442183 AV442183
C 273	17	1.5	296	10	BI051370	BI051370	346	17	1.5	401	9	BE155513	BE155513 PM1-HT035
C 274	17	1.5	297	9	BB282453	BB282453	347	17	1.5	402	10	BE481859	BE481859 167459 BA
C 275	17	1.5	297	12	AQ396241	AQ396241	348	17	1.5	404	9	AI103530	AI103530 EST212819
C 276	17	1.5	298	10	BF654372	BF654372	349	17	1.5	404	10	BF961960	BF961960 QV2-BN004
C 277	17	1.5	299	4	BB714885	Mus_muscu	350	17	1.5	404	10	BF397186	BF397186 UI-R-BN2-
C 278	17	1.5	300	9	AV183288	AV183288	351	17	1.5	405	10	N43877	N43877 yY43f09.r1
C 279	17	1.5	300	9	AV333437	AV333437	352	17	1.5	406	9	AA031840	AA031840 zK17a03.r
C 280	17	1.5	300	9	BB103979	BB103979	353	17	1.5	406	9	AM259327	AM259327 um92a12.y
C 281	17	1.5	300	10	C33963	C33963	354	17	1.5	406	10	BM362465	BM362465 BE53200450
C 282	17	1.5	301	12	A0796342	A0796342	355	17	1.5	408	9	AA815125	AA815125 oc03e07.s
C 283	17	1.5	301	12	BE822824	CMT00019A	356	17	1.5	408	10	H63818	H63818 yT52e06.r1
C 284	17	1.5	305	10	BB530362	BB530362	357	17	1.5	408	12	AQ0563197	AQ0563197 HS_-5302_A
C 285	17	1.5	307	9	BB530362	BB530362	358	17	1.5	409	9	AI845514	AI845514 UI-M-AOI-
C 286	17	1.5	308	9	AA279061	AA279061	359	17	1.5	409	9	BB846348	BB846348 BB846348
C 287	17	1.5	308	9	BB103652	BB103652	360	17	1.5	412	9	AI758015	AI758015 ESTEte06
C 288	17	1.5	309	9	AI109751	AI109751	361	17	1.5	412	9	BE155119	BE155119 PM1-HT035
C 289	17	1.5	309	9	BB098748	BB098748	362	17	1.5	412	12	AQ277745	AQ277745 CITBI-EI-
C 290	17	1.5	310	9	AA443706	AA443706	363	17	1.5	413	9	AA721835	AA721835 33141 Lam
C 291	17	1.5	314	9	AA273869	AA273869	364	17	1.5	413	9	BB800943	BB800943 BB800943
C 292	17	1.5	316	9	AA798776	AA798776	365	17	1.5	415	12	A2762260	A2762260 1M0557105
C 293	17	1.5	322	10	BG642362	BG642362	366	17	1.5	416	9	BB726873	BB726873 BB726873
C 294	17	1.5	324	9	AA916439	AA916439	367	17	1.5	416	12	A2948878	A2948878 2M0212A07
C 295	17	1.5	328	10	D23882	D23882	368	17	1.5	419	9	AI098069	AI098069 uc28b05.x
C 296	17	1.5	329	9	BB259533	BB259533	369	17	1.5	420	9	BB677405	BB677405 BB677405
C 297	17	1.5	335	9	AA259857	umf9d04.y	370	17	1.5	421	10	R91142	R91142 yp94c07.s1
C 298	17	1.5	336	10	BF150920	BF150920	371	17	1.5	421	12	AO613482	AO613482 HS_-5117_B
C 299	17	1.5	340	10	BG053436	RH12_8_H	372	17	1.5	423	9	BE148688	BE148688 MR0-HT024
C 300	17	1.5	342	9	AV543320	AV543320	373	17	1.5	423	12	AQ363157	AQ363157 npxb0055C
C 301	17	1.5	344	12	A2295518	RPCT-23-1	374	17	1.5	425	9	AM143640	AM143640 EST929395
C 302	17	1.5	345	9	AM015489	AM015489	375	17	1.5	425	9	AM919534	AM919534 EST350838
C 303	17	1.5	345	9	AM061528	AM061528	376	17	1.5	425	10	BG629869	BG629869 CC-esf1cL
C 304	17	1.5	345	10	H54475	Yg89C06..r1	377	17	1.5	426	10	BE722840	BE722840 191237 MA
C 305	17	1.5	346	9	AA095474	AA095474	378	17	1.5	426	12	BH224002	BH224002 1006116C0
C 306	17	1.5	351	9	AJ280381	AA3A-AAO-	379	17	1.5	427	12	AO110563	AO110563 CIT-HSP-2
C 307	17	1.5	353	10	D60612	HUM120B08A	380	17	1.5	428	9	AO771689	AO771689 HS_-5412_A
C 308	17	1.5	360	9	AA690738	vu57b03.r	381	17	1.5	430	9	AM502823	AM502823 UI-HF-BNO
C 309	17	1.5	360	10	C40260	C40260	382	17	1.5	430	10	BF738227	BF738227 CM3-KT003

C 383	17	1.5	431	10	BM116916	BM116916 L0843B08-	456	17	1.5	489	9	AI558774	AI558774 fB06001.Y
C 384	17	1.5	431	10	BF229454	BF229454 RC0-C1003	C 457	17	1.5	491	10	BM330902	BM330902 PIC1_67.A
C 385	17	1.5	431	10	BF662363	BF662363 mea84C02.	C 458	17	1.5	492	10	BJ203452	BJ203452 BJ203452
C 386	17	1.5	432	9	BB800416	BB800416 BB800416	C 459	17	1.5	493	10	BC552737	BC552737 dab75C05.
C 387	17	1.5	432	10	BF327099	BF327099 PWO-BN006	C 460	17	1.5	494	9	AI787387	AI787387 uJ31C10.Y
C 388	17	1.5	435	9	BB778342	BB778342 BB778342	C 461	17	1.5	495	9	AA154649	AA154649 m65B12.Y
C 389	17	1.5	436	12	AZ838317	AZ838317 2M0133P22	C 462	17	1.5	496	10	BJ194052	BJ194052 BJ194052
C 390	17	1.5	436	12	AQ040337	AQ040337 HS_5069_B	C 463	17	1.5	496	10	BJ194993	BJ194993 BJ194993
C 391	17	1.5	437	9	BB847450	BB847450 BB847450	C 464	17	1.5	497	12	CNS01N6P	AL151842 Anopheles
C 392	17	1.5	437	10	BF564211	BF564211 UI-R-C4-A	C 465	17	1.5	498	9	AA013912	AA013912 mB24B04.F
C 393	17	1.5	439	9	AM680793	AM680793 WSI_7_E04	C 466	17	1.5	500	9	AU086883	AU086883 AU086883
C 394	17	1.5	439	9	BB684576	BB684576 BB684576	C 467	17	1.5	500	9	AW565686	AW565686 Ld1_348.D
C 395	17	1.5	441	9	BB845915	BB845915 BB845915	C 468	17	1.5	500	10	BE976379	BE976379 BS50A01.Y
C 396	17	1.5	441	10	BI360203	BI360203 387017.MA	C 469	17	1.5	501	10	BE839644	BE839644 RC3-FN015
C 397	17	1.5	443	9	AM563969	AM563969 LG1_275.C	C 470	17	1.5	503	10	BI472831	BI472831 fR92F05.Y
C 398	17	1.5	444	12	AO177112	AO177112 HS_3169.A	C 471	17	1.5	503	10	BM427137	BM427137 pGf2n.PK0
C 399	17	1.5	446	9	AI096270	AI096270 SMAACAC24	C 472	17	1.5	504	10	BE391573	BE391573 UI-R-CA1-
C 400	17	1.5	446	10	BF662111	BF662111 mea84C02.	C 473	17	1.5	504	10	BF602358	BF602358 267796.MA
C 401	17	1.5	447	12	AQ853826	AQ853826 dBxb00C39E	C 474	17	1.5	506	9	AI956311	AI956311 uI73f11.Y
C 402	17	1.5	450	10	BE582699	BE582699 3-1D-MY.P	C 475	17	1.5	506	9	AV672961	AV672961 AV672961
C 403	17	1.5	451	12	AO057114	AO057114 CIT-HSP-2	C 476	17	1.5	507	9	AA987008	AA987008 uc8B1D0.Y
C 404	17	1.5	451	12	AO307457	AO307457 HS_2236.A	C 477	17	1.5	507	10	BF423450	BF423450 sr95f01.Y
C 405	17	1.5	453	12	AZ239057	AZ239057 RPTC1-23-7	C 478	17	1.5	509	9	AA567106	AA567106 CM01093.5
C 406	17	1.5	454	9	AA693862	AA693862 z154f06.s	C 479	17	1.5	509	10	BF016428	BF016428 uy41B05.Y
C 407	17	1.5	454	10	H98235	H98235 YX09d04.s1	C 480	17	1.5	509	12	AQ498953	AQ498953 HS_5154_B
C 408	17	1.5	455	10	BF654326	BF654326 278445.MA	C 481	17	1.5	510	9	AA532220	AA532220 UI-R-C4-a
C 409	17	1.5	456	9	AM271943	AM271943 xs12a07.x	C 482	17	1.5	513	9	AA918328	AA918328 o142d11.s
C 410	17	1.5	457	9	AV673844	AV673844 AV673844	C 483	17	1.5	513	12	AO328654	AO328654 nbxR0043H
C 411	17	1.5	457	9	BE133829	BE133829 ug11905.x	C 484	17	1.5	515	9	AV397832	AV397832 AV397832
C 412	17	1.5	457	10	BE722839	BE722839 191236.MA	C 485	17	1.5	516	12	AO648437	AO648437 RPTC193-DP
C 413	17	1.5	459	12	BH049443	BH049443 RPTC1-24-2	C 486	17	1.5	517	10	BI244541	BI244541 Eg_PSSL0
C 414	17	1.5	460	9	AL035961	AL035961 DKRP2P64E	C 487	17	1.5	519	10	BI244135	BI244135 Eg_CWSL0
C 415	17	1.5	460	10	BI244524	BI244524 Eg_PSSL0	C 488	17	1.5	519	10	BM166200	BM166200 EST568723
C 416	17	1.5	460	12	AZ308431	AZ308431 LM0011G21	C 489	17	1.5	520	10	BI244178	BI244178 Eg_CWSL0
C 417	17	1.5	460	12	AO512667	AO512667 HS_5122.B	C 490	17	1.5	521	10	BG656416	BG656416 jB37B07.X
C 418	17	1.5	462	9	AM721325	AM721325 833013B06	C 491	17	1.5	521	10	W94499	W94499 z613d03.r1
C 419	17	1.5	462	9	AM907652	AM907652 ur96d06.X	C 492	17	1.5	522	9	BB384048	BB384048 BB384048
C 420	17	1.5	462	12	BH283751	BH283751 CH230-49L	C 493	17	1.5	522	12	AQ503934	AQ503934 RPTC1-11-3
C 421	17	1.5	464	9	AM258820	AM258820 um74B02.Y	C 494	17	1.5	523	9	AA848829	AA848829 EST191590
C 422	17	1.5	464	10	AW744632	AW744632 uc13e06.x	C 495	17	1.5	523	10	AM203125	AM203125 f196d12.X
C 423	17	1.5	464	9	BI796941	BI796941 H059E04.E	C 496	17	1.5	523	12	AO282730	AO282730 RPTC11-77
C 424	17	1.5	467	9	AM988801	AM988801 ug11905.Y	C 497	17	1.5	524	9	AW546572	AW546572 L0010B01-
C 425	17	1.5	467	10	BI571748	BI571748 RH07033.5	C 498	17	1.5	524	10	BI244373	BI244373 Eg_PSSL-
C 426	17	1.5	468	9	AA250100	AA250100 mx21d08.T	C 499	17	1.5	524	12	B96297	B96297 T23B22TR.T
C 427	17	1.5	468	9	AM822512	AM822512 ug22b08.X	C 500	17	1.5	525	10	BF6583070	BF6583070 602097960
C 428	17	1.5	470	9	AJ280409	AJ280409 4A3A-AAP-	C 501	17	1.5	526	10	BG5656111	BG5656111 jB37B07.Y
C 429	17	1.5	472	9	AU086871	AU086871 AU086871	C 502	17	1.5	527	9	AV071548	AV071548 EST503060
C 430	17	1.5	472	12	AZ303162	AZ303162 GSSRrC210	C 503	17	1.5	529	10	BG603970	BG603970 EST503060
C 431	17	1.5	473	9	AM023468	AM023468 dF55c11.Y	C 504	17	1.5	529	10	BI066680	BI066680 pGf1n.PK0
C 432	17	1.5	473	12	AZ588503	AZ588503 LM0396G22	C 505	17	1.5	529	12	AO181919	AO181919 HS_3230_A
C 433	17	1.5	474	10	BI371238	BI371238 RE58089.5	C 506	17	1.5	529	12	B52583	B52583 CIT-HSP-200
C 434	17	1.5	475	10	BF545292	BF545292 UI-R-C0-f	C 507	17	1.5	532	9	AV937054	AV937054 C0920G09-
C 435	17	1.5	476	9	AA432637	AA432637 vd90c03.r	C 508	17	1.5	532	10	BM219149	BM219149 C0920G09-
C 436	17	1.5	477	9	BE027344	BE027344 EFBSTea90	C 509	17	1.5	532	12	AO716330	AO716330 HS_5462_A
C 437	17	1.5	478	10	BF606700	BF606700 273813.MA	C 510	17	1.5	534	12	AZ289131	AZ289131 RPTC1-23-1
C 438	17	1.5	479	9	AA711360	AA711360 vt76a06.T	C 511	17	1.5	535	9	AW568179	AW568179 s157e03.Y
C 439	17	1.5	479	9	AJ393592	AJ393592 AJ393592	C 512	17	1.5	536	10	BI240143	BI240143 RE36793.5
C 440	17	1.5	479	10	BG293485	BG293485 602390410	C 513	17	1.5	536	12	AZ892531	AZ892531 RPTC1-24-1
C 441	17	1.5	479	10	BF553118	BF553118 UI-R-C2-n	C 514	17	1.5	537	12	AZ043355	AZ043355 RPTC1-23-3
C 442	17	1.5	480	9	AM982127	AM982127 SSS0047.S	C 515	17	1.5	538	10	BI244103	BI244103 Eg_CWSL0
C 443	17	1.5	480	10	BM352267	BM352267 1984f04.Y	C 516	17	1.5	539	10	BI244391	BI244391 Eg_CWSL-
C 444	17	1.5	480	10	BF518398	BF518398 SSS0156.S	C 517	17	1.5	540	12	AZ739204	AZ739204 RPTC1-24-1
C 445	17	1.5	481	10	BI591176	BI591176 RH04905.5	C 518	17	1.5	542	10	BG664362	BG664362 602798242
C 446	17	1.5	481	10	BE585926	BE585926 Est4PRT-	C 519	17	1.5	543	10	BI983593	BI983593 f85e10.x
C 447	17	1.5	481	12	AZ170649	AZ170649 SP_0117.A	C 520	17	1.5	543	12	AO577215	AO577215 nbxR0090J
C 448	17	1.5	483	12	BH105982	BH105982 RPTC1-24-3	C 521	17	1.5	544	9	AI387230	AI387230 GH17692.5
C 449	17	1.5	484	9	AW318956	AW318956 uni10a03.Y	C 522	17	1.5	545	9	AM055772	AM055772 SMAACAC38
C 450	17	1.5	484	12	AZ288012	AZ288012 RPTC1-23-1	C 523	17	1.5	545	10	BF223660	BF223660 7q78e06.x
C 451	17	1.5	485	9	AI987595	AI987595 uI86d11.Y	C 524	17	1.5	546	10	BM243763	BM243763 KO701E02-
C 452	17	1.5	485	10	BI033926	BI033926 OV2-NN200	C 525	17	1.5	546	12	AO537931	AO537931 RPTC1-11-3
C 453	17	1.5	485	10	N31106	N31106 YX52B01.r1	C 526	17	1.5	547	9	AO447635	AO447635 mgxR0008M
C 454	17	1.5	487	10	BF660083	BF660083 OV2-NN004	C 527	17	1.5	548	9	AI135744	AI135744 GH13606.5
C 455	17	1.5	487	10	BG710769	BG710769 pG11n.PK0	C 528	17	1.5	548	10	BI593456	BI593456 RH1514.5

529	17	1.5	548	12	FR0029414	AL025785 Fugu rubr	c 602	17	1.5	619	12	FR0024450	AL017297 F. rubrip
c 530	17	1.5	549	10	BG521218	BG521218 ps44a05.y	c 603	17	1.5	622	10	BI204276	BI204276 EST522316
c 531	17	1.5	549	10	BI726158	BI726158 1031083E0	c 604	17	1.5	622	10	BI207143	BI207143 EST525183
c 532	17	1.5	550	12	BG008853	BG008853 NM3-GN019	c 605	17	1.5	622	12	AQ258263	AQ258263 nxbx0020A
c 533	17	1.5	550	12	BH210719	Sm1-49P24	c 606	17	1.5	623	12	AV680532	AV680532 AV680532
c 534	17	1.5	551	12	AZ635259	AZ635259 1M0491E04	c 607	17	1.5	623	9	AV865195	AV865195 AV865195
c 535	17	1.5	552	12	AQ451193	AQ451193 HS 5178-A	c 608	17	1.5	623	12	AZ421660	AZ421660 1M0200G01
c 536	17	1.5	556	10	BG082114	BG082114 H3072E05-	c 609	17	1.5	624	10	BF239203	BF239203 053PBE10
c 537	17	1.5	556	10	BJ192507	BJ192507 BI192507	c 610	17	1.5	625	12	AQ292140	AQ292140 nxbx0042B
c 538	17	1.5	556	10	BM079000	BM079000 MEST88-A0	c 611	17	1.5	626	10	AW914666	AW914666 EST345970
c 539	17	1.5	556	10	BM241379	BM241379 BI061B04-	c 612	17	1.5	627	10	BM272063	BM272063 1338E10.Y
c 540	17	1.5	556	12	AQ795506	AQ795506 nxbx0056K	c 613	17	1.5	629	10	BG972168	BG972168 602840929
c 541	17	1.5	559	9	BE232713	BE232713 137735 MA	c 614	17	1.5	627	10	BE356820	BE356820 DC1_12_D0
c 542	17	1.5	560	9	AI405680	AI405680 GH5782.3	c 615	17	1.5	627	12	AQ258263	AQ258263 nxbx0037L
c 543	17	1.5	560	10	BG914823	BG914823 602813636	c 616	17	1.5	628	9	BB660733	BB660733 BB660733
c 544	17	1.5	562	12	BH383454	Ag-ND-124	c 617	17	1.5	629	9	AI063698	AI063698 GH03654.5
c 545	17	1.5	563	12	AC687886	AC687886 nxbx0075L	c 618	17	1.5	629	9	BB660766	BB660766 BB660766
c 546	17	1.5	563	12	AZ762340	AZ762340 1M0557P16	c 619	17	1.5	630	10	BJ123708	BJ123708 BJ123708
c 547	17	1.5	566	9	BE107198	BE107198 UT-R-BS1-	c 620	17	1.5	630	12	BH393538	BH393538 AC-ND-169
c 548	17	1.5	567	10	BI370353	BI370353 RES6842.5	c 621	17	1.5	631	9	BB661253	BB661253 BB661253
c 549	17	1.5	568	9	AI314326	AI314326 uJ29f07.x	c 622	17	1.5	633	9	AU167237	AU167237 AU167237
c 550	17	1.5	568	12	AQ510299	AQ510299 nxbx0095M	c 623	17	1.5	633	12	BH296781	BH296781 CH230-187
c 551	17	1.5	570	9	AV567175	AV567175 AV567175	c 624	17	1.5	634	12	AZ519933	AZ519933 RPCI-11-2
c 552	17	1.5	570	10	BM306713	BM306713 REESTee34	c 625	17	1.5	639	9	AA458239	AA458239 vq46h06.r
c 553	17	1.5	571	10	BM243488	BM243488 RM647E08-	c 626	17	1.5	640	12	AZ861355	AZ861355 2M0167F17
c 554	17	1.5	573	9	AI558562	AI558562 fb3f112.y	c 627	17	1.5	643	10	BI082900	BI082900 602878263
c 555	17	1.5	573	10	BI344134	BI344134 372515 MA	c 628	17	1.5	643	12	AZ958789	AZ958789 2M0226A18
c 556	17	1.5	574	12	PT003C11R	AL447527 Parmeciu	c 629	17	1.5	644	10	BF780746	BF780746 602103019
c 557	17	1.5	574	10	BG970165	BG970165 602839310	c 630	17	1.5	645	10	BI609052	BI609052 RH13785.5
c 558	17	1.5	576	12	AO666706	HS 2108-A	c 631	17	1.5	645	10	BE260304	BE260304 601151154
c 559	17	1.5	576	10	BF793207	602253054	c 632	17	1.5	645	12	AZ136272	AZ136272 SP-0172_A
c 560	17	1.5	576	10	BI232061	BI232061 RE28067.5	c 633	17	1.5	645	12	AZ375406	AZ375406 1M0128C1A
c 561	17	1.5	577	10	BI344131	BI344131 372509 MA	c 634	17	1.5	645	12	BH345789	BH345789 CH230-52D
c 562	17	1.5	578	9	AW259726	AW259726 umf5h04.y	c 635	17	1.5	646	10	BI1997341	BI1997341 103104990
c 563	17	1.5	580	12	AQ746965	HS 5538-A	c 636	17	1.5	646	10	BE309600	BE309600 601094867
c 564	17	1.5	581	10	BG174482	BG174482 602334310	c 637	17	1.5	647	10	BM426032	BM426032 pgf2a.pk0
c 565	17	1.5	581	10	BG916215	BG916215 602814930	c 638	17	1.5	648	10	BI171494	BI171494 RE13283.5
c 566	17	1.5	581	10	BI571497	BI571497 RM06630.5	c 639	17	1.5	648	10	BM426349	BM426349 pgf2n.pk0
c 567	17	1.5	582	10	BG865021	BG865021 602799489	c 640	17	1.5	649	10	BI112112	BI112112 602899769
c 568	17	1.5	582	12	BH105484	BH105484 RPCI-24-3	c 641	17	1.5	650	10	BM440229	BM440229 pgf1n.pk0
c 569	17	1.5	583	10	BI357357	BI357357 RE43822.5	c 642	17	1.5	651	9	AL565499	AL565499 AL565499
c 570	17	1.5	584	10	BI747987	BI747987 rm48408.y	c 643	17	1.5	651	12	AO576207	AO576207 nxbx0088D
c 571	17	1.5	586	9	AI116773	AI116773 mq24B05.r	c 644	17	1.5	653	9	AI786749	AI786749 UK53e04.x
c 572	17	1.5	586	9	AM108068	AM108068 umf19b10.x	c 645	17	1.5	653	10	BM440091	BM440091 pgf1n.pk0
c 573	17	1.5	586	9	BE148547	BE148547 MRO-HT024	c 646	17	1.5	654	9	AV873699	AV873699 AV873699
c 574	17	1.5	587	9	AI138111	AI138111 mq30Q03.r	c 647	17	1.5	654	10	BE306066	BE306066 601101772
c 575	17	1.5	591	10	BI181441	BI181441 BI181441	c 648	17	1.5	654	12	AZ522336	AZ522336 205PDE06
c 576	17	1.5	591	12	BH073306	BH073306 RPCI-24-2	c 649	17	1.5	654	12	AG105742	AG105742 Pan trogl
c 577	17	1.5	592	9	AI135769	AI135769 GH13637.5	c 650	17	1.5	654	10	BI357656	BI357656 RE44186.5
c 578	17	1.5	595	10	BI903723	BI903723 603167346	c 651	17	1.5	656	10	BI508032	BI508032 BB170006B
c 579	17	1.5	595	10	BM450914	BM450914 PGP2n.pk0	c 652	17	1.5	656	12	AQ287743	AQ287743 nxbx0031C
c 580	17	1.5	595	12	AZ968226	AZ968226 2M0240B05	c 653	17	1.5	656	10	BF703659	BF703659 MI-P-E4-a
c 581	17	1.5	596	10	BI153636	BI153636 602871218	c 654	17	1.5	660	12	AO156977	AO156977 nxbx0008B
c 582	17	1.5	596	10	BI432494	BI432494 EST535255	c 655	17	1.5	660	12	AG170224	AG170224 Pan trogl
c 583	17	1.5	601	10	BI720349	BI720349 1031049C0	c 656	17	1.5	661	9	AM611375	AM611375 um64B02.x
c 584	17	1.5	602	10	BG687509	BG687509 602788440	c 657	17	1.5	662	9	BE213448	BE213448 GE-FV-P6B
c 585	17	1.5	602	10	BG935061	SL1-0367	c 658	17	1.5	662	12	AZ115588	AZ115588 RPCI-23-4
c 586	17	1.5	605	12	AZ384117	AZ384117 1M0142G03	c 659	17	1.5	664	12	AG067369	AG067369 Pan trogl
c 587	17	1.5	605	12	AQ256082	AQ256082 nxbx0015I	c 660	17	1.5	665	12	AG146741	AG146741 Pan trogl
c 588	17	1.5	606	10	BI171492	BI171492 RE13281.5	c 661	17	1.5	667	12	BH154248	BH154248 EMTSO03TF
c 589	17	1.5	607	10	BF780576	BF780576 602103119	c 662	17	1.5	667	12	BH211577	BH211577 SM1-62P22
c 590	17	1.5	608	9	AW284185	AW284185 uL1-262-G	c 663	17	1.5	668	10	BI367114	BI367114 RE52802.5
c 591	17	1.5	608	9	AW912584	AW912584 uL51f05.x	c 664	17	1.5	668	12	AZ524151	AZ524151 227PBH03
c 592	17	1.5	610	12	AQ326265	AQ326265 nxbx0028K	c 665	17	1.5	669	9	AM260329	AM260329 umf76908.x
c 593	17	1.5	610	12	AI794545	AI794545 fc45f10.y	c 666	17	1.5	669	10	BE250819	BE250819 601106826
c 594	17	1.5	612	10	AZ110018	AZ110018 RPCI-23-4	c 667	17	1.5	671	10	BE381936	BE381936 601272476
c 595	17	1.5	612	10	BI513089	BI513089 RB160011A	c 668	17	1.5	672	9	AM260327	AM260327 umf6902.x
c 596	17	1.5	612	10	BI616540	BI616540 RH46517.5	c 669	17	1.5	673	12	BH523324	BH523324 BCG1H88TR
c 597	17	1.5	612	12	AZ838260	AZ838260 2M0133F20	c 670	17	1.5	674	9	AI267144	AI267144 uJ11d10.x
c 598	17	1.5	613	12	AA142182	AA142182 CK0033.3	c 671	17	1.5	674	12	AZ838319	AZ838319 2M0133P24
c 599	17	1.5	616	9	AI385924	AI385924 mq24B05.y	c 672	17	1.5	676	9	AI517824	AI517824 GH28888.5
c 600	17	1.5	619	12	AZ379988	AZ379988 1M0135G21	c 673	17	1.5	678	12	AG069788	AG069788 Pan trogl
c 601	17	1.5	619	12	AZ381452	AZ381452 1M0138J02	c 674	17	1.5	679	9	BB305534	BB305534 BB305534

C 675	17	1.5	679	10	BE532770	BE532770 601233916	C 748	17	1.5	762	9	AV404062
C 676	17	1.5	679	12	AQ566429	AQ566429 HS-2096.A	C 749	17	1.5	763	10	B1558679
C 677	17	1.5	681	10	B1193392	B1193392 602947024	C 750	17	1.5	764	9	A1137374
C 678	17	1.5	682	12	AZ524027	AZ524027 226PBC10	C 751	17	1.5	767	12	BH399000
C 679	17	1.5	683	10	AW913495	AW913495 uf53f05.y	C 752	17	1.5	768	10	BG972346
C 680	17	1.5	683	10	B1102350	B1102350 602885691	C 753	17	1.5	769	9	AU003365
C 681	17	1.5	684	10	AW319624	AW319624 um99f06.x	C 754	17	1.5	769	12	BH510892
C 682	17	1.5	685	9	A1514098	A1514098 u112b10.x	C 755	17	1.5	772	12	AQ739758
C 683	17	1.5	686	9	A1513977	A1513977 GH27011.5	C 756	17	1.5	780	9	AA968100
C 684	17	1.5	688	9	AU004500	AU004500 AU004500	C 757	17	1.5	780	10	BG422548
C 685	17	1.5	690	10	AM261689	AM261689 um08h01.x	C 758	17	1.5	781	9	AU122743
C 686	17	1.5	690	10	B1213754	B1213754 RE19247.5	C 759	17	1.5	781	10	B1102803
C 687	17	1.5	692	9	BE148614	BE148614 MR0-HR024	C 760	17	1.5	782	10	BF687125
C 688	17	1.5	692	12	AZ184785	AZ184785 SP-1003.B	C 761	17	1.5	783	12	AQ857998
C 689	17	1.5	693	9	A1267069	A1267069 u110d08.x	C 762	17	1.5	784	9	A1115991
C 690	17	1.5	694	9	AU138198	AU138198 AU138198	C 763	17	1.5	784	12	AO876204
C 691	17	1.5	696	9	AM260285	AM260285 um76a11.x	C 764	17	1.5	788	12	BH559362
C 692	17	1.5	698	9	AW319673	AW319673 um01h10.x	C 765	17	1.5	790	9	AU079836
C 693	17	1.5	698	12	AQ256136	AQ256136 nbxb0011D	C 766	17	1.5	790	10	B1143892
C 694	17	1.5	699	9	A1649039	A1649039 uk34e04.x	C 767	17	1.5	791	12	AO687266
C 695	17	1.5	699	12	AQ075055	AQ075055 CIT-HSP-2	C 768	17	1.5	791	12	AZ202357
C 696	17	1.5	699	12	BH456484	BH456484 BOHRT54TF	C 769	17	1.5	791	12	BH446260
C 697	17	1.5	702	9	A1267142	A1267142 u11d08.x	C 770	17	1.5	791	12	AQ292006
C 698	17	1.5	702	9	AU004414	AU004414 AU004414	C 771	17	1.5	791	12	CNS03080
C 699	17	1.5	704	12	AQ529553	AQ529553 RPCI-11-3	C 772	17	1.5	795	12	CNS0302L
C 700	17	1.5	705	12	AL654122	AL654122 AL654122	C 773	17	1.5	796	9	BF787922
C 701	17	1.5	705	12	BH183124	BH183124 022.L-18-	C 774	17	1.5	797	12	BF787922
C 702	17	1.5	708	12	CNS070D5	AL650075.T3 end of	C 775	17	1.5	797	12	BG707213
C 703	17	1.5	708	9	A1132447	A1132447 ue25c10.x	C 776	17	1.5	800	9	AL119083
C 704	17	1.5	708	9	AL556141	AL556141 AL556141	C 777	17	1.5	800	9	AU079850
C 705	17	1.5	709	9	A1118368	A1118368 ue23c04.x	C 778	17	1.5	803	12	AQ575528
C 706	17	1.5	711	9	A1649353	A1649353 uk30905.x	C 779	17	1.5	808	10	B1664288
C 707	17	1.5	715	9	A1405844	A1405844 GH25985.5	C 780	17	1.5	809	12	AQ257607
C 708	17	1.5	716	9	BJ088409	BJ088409 BJ088409	C 781	17	1.5	810	10	B1117268
C 709	17	1.5	717	10	BG471712	BG471712 602513436	C 782	17	1.5	810	10	B1664467
C 710	17	1.5	718	12	AZ660690	AZ660690 1M0538372	C 783	17	1.5	810	12	CNS00920
C 711	17	1.5	720	9	AU136782	AU136782 AU136782	C 784	17	1.5	811	9	AW107500
C 712	17	1.5	720	10	BE314237	BE314237 601153044	C 785	17	1.5	811	12	BH382329
C 713	17	1.5	720	10	BF320332	BF320332 u252f04.y	C 786	17	1.5	815	9	AA968324
C 714	17	1.5	721	9	A1316510	A1316510 u114f02.x	C 787	17	1.5	816	9	AW106903
C 715	17	1.5	721	9	A1597338	A1597338 u190a08.y	C 788	17	1.5	817	12	BH095860
C 716	17	1.5	721	12	AQ576521	AQ576521 nbxb0089E	C 789	17	1.5	820	9	A1788951
C 717	17	1.5	722	9	A1132433	A1132433 ue25b07.x	C 790	17	1.5	820	10	BF788778
C 718	17	1.5	722	10	BF788991	BF788991 602104835	C 791	17	1.5	821	10	B1218037
C 719	17	1.5	723	12	AQ687207	AQ687207 nbxb0073P	C 792	17	1.5	823	12	AO686566
C 720	17	1.5	723	12	BH395453	BH395453 AG-ND-147	C 793	17	1.5	823	12	BH568797
C 721	17	1.5	724	10	BE914814	BE914814 601667838	C 794	17	1.5	824	10	B1R54319
C 722	17	1.5	724	12	BH300162	BH300162 CH230-109	C 795	17	1.5	826	9	A1790678
C 723	17	1.5	725	9	A1314702	A1314702 u127g01.x	C 796	17	1.5	828	10	BG567790
C 724	17	1.5	725	10	BG718515	BG718515 602669573	C 797	17	1.5	829	9	A1649179
C 725	17	1.5	725	10	B1103054	B1103054 602888689	C 798	17	1.5	833	9	A1647699
C 726	17	1.5	725	12	AZ985759	AZ985759 2M0267P12	C 799	17	1.5	833	10	B1554921
C 727	17	1.5	726	12	AZ131936	AZ131936 CSUNB011	C 800	17	1.5	836	10	BE268283
C 728	17	1.5	726	12	AZ613041	AZ613041 1M0441P12	C 801	17	1.5	843	9	A1648972
C 729	17	1.5	727	10	BE284099	BE284099 601099296	C 802	17	1.5	845	12	AZ532784
C 730	17	1.5	727	10	BE306118	BE306118 601100414	C 803	17	1.5	846	12	BH562524
C 731	17	1.5	728	12	AQ289732	AQ289732 nbxb0035H	C 804	17	1.5	848	9	AW106932
C 732	17	1.5	734	10	BF584963	BF584963 602090849	C 805	17	1.5	851	10	BG495516
C 733	17	1.5	735	10	B1412195	B1412195 602989986	C 806	17	1.5	854	12	AO687775
C 734	17	1.5	737	12	AG097007	AG097007 Pan Trogl	C 807	17	1.5	857	9	A1746298
C 735	17	1.5	738	9	AV700539	AV700539 AV700539	C 808	17	1.5	857	10	BF786310
C 736	17	1.5	743	10	BG809267	BG809267 m9he0018x	C 809	17	1.5	857	12	BH092908
C 737	17	1.5	743	10	BG969294	BG969294 602836765	C 810	17	1.5	861	9	A1113964
C 738	17	1.5	746	12	AZ879307	AZ879307 RPCI-23-1	C 811	17	1.5	861	9	A1987711
C 739	17	1.5	747	12	BH581640	BH581640 BOHLEF30TF	C 812	17	1.5	862	10	B1663369
C 740	17	1.5	747	12	CNS00BE7	CNS00BE7 CROSOPI11	C 813	17	1.5	862	10	B1948892
C 741	17	1.5	750	12	AQ349409	AQ349409 RPCI11-11	C 814	17	1.5	864	12	B1260075
C 742	17	1.5	751	9	AU167238	AU167238 AU167238	C 815	17	1.5	864	12	CNS00913
C 743	17	1.5	753	10	B1154720	B1154720 602904512	C 816	17	1.5	867	10	BH874924
C 744	17	1.5	754	12	AQ578259	AQ578259 nbxb0092F	C 817	17	1.5	868	10	BG300655
C 745	17	1.5	756	12	AQ326081	AQ326081 nbxb00024L	C 818	17	1.5	869	10	B1143894
C 746	17	1.5	757	10	BG765608	BG765608 602738502	C 819	17	1.5	870	9	AL537110
C 747	17	1.5	760	9	AW475206	AW475206 um72h10.y	C 820	17	1.5	871	10	B1R69842

C 821	17	1.5	875	9	AA203353	AA203353	zx57b09..r	C 894	17	1.5	1101	12	CNS002V9	AL063511	Drosophila
C 822	17	1.5	876	10	B1412109	B1412109	602964860	C 895	17	1.5	1101	12	CNS00CSN	AL059688	Drosophila
C 823	17	1.5	876	10	B1646433	B1646433	603279010	C 896	17	1.5	1101	12	CNS00HVZ	AL074113	Drosophila
C 824	17	1.5	879	12	CNS050MS	AL345493	Tetradon	C 897	17	1.5	1101	12	CNS05471	AL032091	Tetradon
C 825	17	1.5	881	10	B1147088	B1147088	602913360	C 898	17	1.5	1102	10	EM460639	B4460639	AGERCORT
C 826	17	1.5	884	10	BG295757	BG295757	602393210	C 899	17	1.5	1147	10	BG024236	BG024236	602302954
C 827	17	1.5	884	10	BES70618	BES70618	601329923	C 900	17	1.5	1156	10	BET29821	BET29821	601564735
C 828	17	1.5	886	10	BE784356	BE784356	602110131	C 901	17	1.5	1160	12	CNS007J9	AL067337	Drosophila
C 829	17	1.5	891	10	BE728760	BE728760	601564447	C 902	17	1.5	1201	12	CNS0010X	AL060977	Drosophila
C 830	17	1.5	892	10	BE792416	BE792416	601585162	C 903	17	1.5	1227	10	BG260063	BG260063	601150949
C 831	17	1.5	892	12	AZ681927	AZ681927	ENTVI11TR	C 904	17	1.5	1321	10	BE166709	BE166709	601776224
C 832	17	1.5	892	12	BH157767	BH157767	ENTR821F	C 905	17	1.5	1339	10	BE294378	BE294378	004PBE01
C 833	17	1.5	893	10	B1410036	B1410036	602962519	C 906	17	1.5	1522	10	BE694004	BE694004	602082592
C 834	17	1.5	897	12	CNS030DM7	AL261304	Tetradon	C 907	17	1.5	1565	11	AK015326	AK015326	Mus muscu
C 835	17	1.5	898	10	BF315074	BF315074	601902510	C 908	17	1.5	1602	10	BG165628	BG165628	602345114
C 836	17	1.5	898	12	AZ691137	AZ691137	ENTM682TF	C 909	17	1.5	1813	10	BE237826	BE237826	601842371
C 837	17	1.5	899	12	BH133763	BH133763	ENTM682TF	C 910	17	1.5	2054	11	BC022902	BC022902	Mus muscu
C 838	17	1.5	900	9	AL666200	AL666200	AL666200	C 911	17	1.5	3065	11	AK019548	AK019548	Mus muscu
C 839	17	1.5	900	12	BH177432	BH177432	009_D_10-	C 912	17	1.4	50	10	BG223087	BG223087	nah13604
C 840	17	1.5	900	12	CNS07J38	AL614390	T7 end of	C 913	16	1.4	55	10	BE542738	BE542738	601066341
C 841	17	1.5	901	12	AZ535919	AZ535919	ENTB096TR	C 914	16	1.4	61	9	AA546711	AA546711	VK55C02..r
C 842	17	1.5	901	12	BH159707	BH159707	ENTTC73TF	C 915	16	1.4	87	12	AZ780134	AZ780134	2M0017108
C 843	17	1.5	903	10	BF785598	BF785598	602112386	C 916	16	1.4	104	9	AU209499	AU209499	AU209499
C 844	17	1.5	904	12	BF785598	BH167215	ENTSO18TR	C 917	16	1.4	106	10	BG893358	BG893358	dae19702
C 845	17	1.5	911	10	BF780753	BF780753	602103027	C 918	16	1.4	111	9	AA591875	AA591875	V147H04..r
C 846	17	1.5	914	10	BF780409	BF780409	602103928	C 919	16	1.4	126	9	AM881548	AM881548	RC4-OT003
C 847	17	1.5	915	10	BG497847	BG497847	602543137	C 920	16	1.4	127	12	AZ310281	AZ310281	1M0025K12
C 848	17	1.5	920	12	CNS03MA1	AL250498	Tetradon	C 921	16	1.4	129	9	BE167709	BE167709	CM3-HT051
C 849	17	1.5	923	10	BH102605	BH102605	602888128	C 922	16	1.4	131	10	B1435853	B1435853	EST538614
C 850	17	1.5	923	12	CNS060UJ	AL415277	T7 end of	C 923	16	1.4	136	10	BE937719	BE937719	PM2-TN002
C 851	17	1.5	927	12	AQ688304	AQ688304	nxb00767K	C 924	16	1.4	137	9	A1508628	A1508628	vb53g11..y
C 852	17	1.5	928	10	BF787496	BF787496	602110467	C 925	16	1.4	143	9	AA386760	AA386760	vb53g11..y
C 853	17	1.5	928	10	BF787496	BF787496	602113914	C 926	16	1.4	143	12	AZ239310	AZ239310	RCPT-23-7
C 854	17	1.5	928	10	BF787496	BF787496	602113914	C 927	16	1.4	142	10	T80393	T80393	y405d12..r1
C 855	17	1.5	929	12	CNS00CDN	AL056744	Drosophila	C 928	16	1.4	146	10	B1002006	B1002006	MR2-HN006
C 856	17	1.5	931	10	BG119272	BG119272	602349145	C 929	16	1.4	147	9	AA069049	AA069049	zm11h02..r
C 857	17	1.5	931	10	BF344941	BF344941	602014259	C 930	16	1.4	148	9	AA880702	AA880702	vx30d02..r
C 858	17	1.5	933	12	CNS0310C	AL245901	Tetradon	C 931	16	1.4	149	9	BE072686	BE072686	RC2-BT054
C 859	17	1.5	934	10	BG473633	BG473633	602515986	C 932	16	1.4	152	10	BF063806	BF063806	7H96d08..x
C 860	17	1.5	934	10	BF311362	BF311362	601896765	C 933	16	1.4	153	10	T81194	T81194	y405d12..r1
C 861	17	1.5	935	12	AZ672871	AZ672871	ENTM657TF	C 934	16	1.4	163	12	AO658652	AO658652	Sheared D
C 862	17	1.5	935	12	BH138870	BH138870	ENTM037TF	C 935	16	1.4	163	9	BE072850	BE072850	RC2-BT054
C 863	17	1.5	938	10	BM36759	BM36759	5009-0-24	C 936	16	1.4	165	9	AM368869	AM368869	IL5-HT019
C 864	17	1.5	940	12	CNS02M62	AL203699	Tetradon	C 937	16	1.4	169	9	AV035412	AV035412	AV035412
C 865	17	1.5	944	10	BF159933	BF159933	601768105	C 938	16	1.4	171	9	A1555058	A1555058	UI-R-C2P-
C 866	17	1.5	947	10	BG259538	BG259538	602378657	C 939	16	1.4	175	9	AM454193	AM454193	zebl1070
C 867	17	1.5	948	9	A1513805	A1513805	GH26904..5	C 940	16	1.4	179	9	BR200792	BR200792	BR200792
C 868	17	1.5	953	12	BH155124	BH155124	ENTM811TR	C 941	16	1.4	181	9	BE072681	BE072681	RC2-BT054
C 869	17	1.5	959	12	CNS05HQ	AL337895	Tetradon	C 942	16	1.4	186	10	B1127896	B1127896	G067P62Y
C 870	17	1.5	960	9	AL666445	AL666445	AL666445	C 943	16	1.4	188	9	AA420191	AA420191	v51f108..r
C 871	17	1.5	964	12	CNS05AYD	AL329134	Tetradon	C 944	16	1.4	188	10	B1169905	B1169905	REI11212..5
C 872	17	1.5	965	10	BM008624	BM008624	603618274	C 945	16	1.4	188	12	BH341484	BH341484	CH230-111
C 873	17	1.5	970	12	CNS02X0J	AL217756	Tetradon	C 946	16	1.4	191	12	BE156433	BE156433	OVO-HT036
C 874	17	1.5	972	12	CNS07AMS	AL436778	T7 end of	C 947	16	1.4	191	10	BF639351	BF639351	NEP011061
C 875	17	1.5	973	10	BI739342	BI739342	603359652	C 948	16	1.4	199	12	AO635215	AO635215	SB84-AGS-
C 876	17	1.5	973	12	CNS03VAV	AL262195	Tetradon	C 949	16	1.4	202	9	AM3332410	AM3332410	AV143483
C 877	17	1.5	974	12	CNS005MF	AL059381	Drosophila	C 950	16	1.4	203	9	AV143483	AV143483	AV143483
C 878	17	1.5	985	12	CNS005M4	AL059370	Drosophila	C 951	16	1.4	206	9	AV359607	AV359607	AV359607
C 879	17	1.5	997	10	BF687782	BF687782	602066832	C 952	16	1.4	207	9	BE162771	BE162771	OV3-HT045
C 880	17	1.5	1001	10	B1414294	B1414294	602986585	C 953	16	1.4	208	9	AM314065	AM314065	9842-MARC
C 881	17	1.5	1002	10	BG755130	BG755130	602711446	C 954	16	1.4	211	12	BH033619	BH033619	RCPT-24-3
C 882	17	1.5	1007	10	BG297012	BG297012	602394862	C 955	16	1.4	214	12	AZ752441	AZ752441	RCPT-24-1
C 883	17	1.5	1021	12	CNS006JY	AL071404	Drosophila	C 956	16	1.4	214	10	BF761364	BF761364	RC2-CS001
C 884	17	1.5	1024	12	CNS0374L	AL259374	Tetradon	C 957	16	1.4	216	9	BH147886	BH147886	B147886
C 885	17	1.5	1030	12	CNS07BGO	AL437854	T7 end of	C 958	16	1.4	216	10	BG346537	BG346537	daa38a02..
C 886	17	1.5	1031	12	BI2303	BI2303	F16B11--T7 I	C 959	16	1.4	217	10	BG156300	BG156300	sa873h10..
C 887	17	1.5	1035	12	CNS04ODT	AL288778	Tetradon	C 960	16	1.4	217	10	BF611291	BF611291	dg77h12..y
C 888	17	1.5	1046	10	BF785777	BF785777	602112294	C 961	16	1.4	220	9	A1950291	A1950291	wq38h05..x
C 889	17	1.5	1049	12	CNS04MPZ	AL297142	Tetradon	C 962	16	1.4	221	9	BE049807	BE049807	NXNY_144-
C 890	17	1.5	1085	12	CNS00ZXP	AL088295	Drosophila	C 963	16	1.4	222	9	AM649477	AM649477	EST327931
C 891	17	1.5	1090	12	CNS05BYF	AL330433	Tetradon	C 964	16	1.4	223	10	BM278729	BM278729	As-fgk..66
C 892	17	1.5	1101	12	CNS000WZ	AL063247	Drosophila	C 965	16	1.4	223	10	H30565	H30565	y44f503..s1
C 893	17	1.5	1101	12	CNS001T6	AL061332	Drosophila	C 966	16	1.4	227	9	AV885030	AV885030	AV885030

C 967	16	1.4	227	10	BG346546	BG346546 daa39a12.
C 968	16	1.4	229	9	BB057250	BB057250 BB057250
C 969	16	1.4	230	10	BF199025	BF199025 249060 MA
C 970	16	1.4	232	9	BB186527	BB186527 BB186527
C 971	16	1.4	233	10	BJ004458	BJ004458 BJ004458
C 972	16	1.4	233	10	H39193	H39193 yn04d1.r1
C 973	16	1.4	233	10	T39491	T39491 ya06f01.r1
C 974	16	1.4	235	9	AV356606	AV356606 AV356606
C 975	16	1.4	235	10	T75805	T75805 10583 Lambda
C 976	16	1.4	236	9	AV217051	AV217051 AV217051
C 977	16	1.4	236	10	BI887991	BI887991 ZF637-1-0
C 978	16	1.4	237	9	AA078111	AA078111 714C03 C
C 979	16	1.4	237	10	F05928	F05928 HSC0RD061 n
C 980	16	1.4	238	9	A1764895	A1764895 MEG0750
C 981	16	1.4	238	9	A1772835	A1772835 EST25935
C 982	16	1.4	238	9	AV336510	AV336510 AV336510
C 983	16	1.4	238	9	AV561972	AV561972 AV561972
C 984	16	1.4	238	10	BG124622	BG124622 EST470184
C 985	16	1.4	240	9	AU073887	AU073887 AU073887
C 986	16	1.4	240	10	BM434261	BM434261 1RT0C12
C 987	16	1.4	241	9	AW843459	AW843459 CM0-CM003
C 988	16	1.4	241	9	BB198000	BB198000 BB198000
C 989	16	1.4	241	10	BE583001	BE583001 3-4H-MY P
C 990	16	1.4	241	12	AZ009581	AZ009581 RPCI-23-3-3
C 991	16	1.4	241	12	AO193779	AO193779 CIT-HSP-2
C 992	16	1.4	242	9	BB808166	BB808166 BB808166
C 993	16	1.4	242	10	BG630074	BG630074 CC-esf1CL
C 994	16	1.4	243	9	AV229167	AV229167 AV229167
C 995	16	1.4	243	12	AZ645100	AZ645100 1M0510B19
C 996	16	1.4	247	10	BI730868	BI730868 603351891
C 997	16	1.4	249	9	AV275280	AV275280 AV275280
C 998	16	1.4	249	9	AV315445	AV315445 AV315445
C 999	16	1.4	252	9	AV230398	AV230398 AV230398
C 1000	16	1.4	252	10	BG613665	BG613665 CC-esf1CL

## ALIGNMENTS

RESULT 1  
 LOCUS A1176234/c 348 bp mRNA linear EST 20-JAN-1999  
 DEFINITION ROVB0921815 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone  
 accession A1176234 end, mRNA sequence.  
 version A1176234  
 keywords A1176234.1 GI:3726872  
 source EST.  
 organism Rattus sp.  
 Rattus sp.  
 Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 348)  
 Lee,N.H., Gloddek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.  
 Rat Genome Project: Generation of a Rat EST (RESTR) Catalog & Rat  
 Gene Index  
 Unpublished (1998)  
 other ESTs: TC48377  
 Contact: lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 Location/Qualifiers  
 source  
 1..348

FEATURES	SOURCE
location/Qualifiers	
1. 348	
/organism="Rattus sp."	
/db_xref="ATCC (inhost):2031231"	
/db_xref="taxon:10118"	
/clone="ROYBO92"	
/clone_1lb="Normalized rat ovary, Bento Soares"	

```

/note="Organ: ovary; Vector: pUT13pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT      79 a      85 c      104 g      80 t
ORIGIN
Query Match      1.9%; Score 22; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
Oy 935 tcatggacttcacacctgatgtt 956
|||||
Db 331 tcatggaccttcacacctgatgtt 310

```

RESULT	2
Locus	A1178895/c
Definition	A1178895 354 bp mRNA linear EST 20-JAN-1998
LOCUS	EST222577 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone RSPBQ78.3 end, mRNA sequence.
Accession	A1178895
Version	A1178895.1 GI:3729533
Keywords	EST,
Source	Rattus sp.
Organism	Rattus sp. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
Reference	1 (bases 1 to 354) Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J., Kerlavage, A.R. and Adams, M.D.
Authors	Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat Gene Index
Title	Unpublished (1998)
Journal	Other_ESTS: TC48377
Comment	

FEATURES		location/Qualifiers
SOURCE	1..354	/organism="Rattus sp." /db_xref="AFCC (lnhost):2034870" /db_xref="taxon:10118" /clone="RSPBQ78" /clone_1lb="Normalized rat spleen, Bento Soares"
		/note="Organ: spleen; Vector: pTV3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT	85 a 86 c 112 g 77 t	
ORIGIN		
Query Match	1.9%; Score 22; DB 9; Length 354;	
Best Local Similarity	100.0%; Pred. No. 2;	
Matches	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY 935	tcatgagcctcacccgtatgtt 956	
Db 310	TCATGGCGCTCACCCGTGTT 289	
RESULT 3		
A0187817		585 bp DNA linear GSS 03-NOV-1998
LOCUS	A0187817	
DEFINITION	HS-3141_B1.D09_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3141 Col=17 Row=H, DNA sequence.	
ACCESSION	A0187817	
VERSION	A0187817.1 GI:3587229	
KEYWORDS	GSS.	
SOURCE	human.	

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 585)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,R., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htsc.washington.edu>

FEATURES  
SOURCE  
Location/Qualifiers  
1..585  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate=3142 Col=17 Row=H"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
E-Coli DH10B"  
BASE COUNT 149 a 135 c 108 g 187 t 6 others  
ORIGIN

Query Match 1.8%; Score 21; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 46 tgctgcctgcctgctactat 66  
|||||  
Db 93 TGCTGCTGCTGCTACTTAT 113

RESULT 4  
LOCUS A0891144 627 bp DNA linear GSS 10-NOV-1999  
DEFINITION HS-3142\_AL-A06\_MR CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate=3142 Col=11 Row=A, DNA sequence.  
ACCESSION A0891144  
VERSION A0891144.1 GI:6347334  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 627)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,R., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 3142 row: A column: 11  
Seq primer: M13 Reverse  
Class: BAC ends  
High quality sequence stop: 627.  
Location/Qualifiers  
1..627  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate=3142 Col=11 Row=A"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
E-Coli DH10B"  
BASE COUNT 154 a 129 c 105 g 221 t 18 others  
ORIGIN

Query Match 1.8%; Score 21; DB 12; Length 627;  
Best Local Similarity 100.0%; Pred. No. 7.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 46 tgctgcctgcctgctactat 66  
|||||  
Db 120 TGCTGCTGCTGCTACTTAT 140

RESULT 5  
LOCUS BB515310/c 282 bp mRNA linear EST 28-JUL-2000  
DEFINITION BB515310 RIKEN full-length enriched, 16 days neonate heart Mus  
musculus cDNA clone D830007M21 3' similar to NM\_008509 Mus musculus  
lipoprotein lipase (LPL), mRNA sequence.  
ACCESSION BB515310  
VERSION BB515310.1 GI:9566768  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 282)  
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Araiwa,T., Carninci,  
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,  
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,  
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,  
Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,  
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomihaga,N., Toya,  
T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,  
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,  
M., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Konno,H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL:<http://genome.gsc.riken.go.jp/>  
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,  
N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Thermolabile and thermocatalytic of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh,M., Katsunari,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,  
Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

FEATURES  
source Location/Qualifiers

1.282  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="D83007M21"  
/clone\_1lb="RIKEN full-length enriched, 16 days neonate  
heart"  
/tissue\_type="heart"  
/dev\_stage="16 days neonate"  
/lab\_host="DH10B"  
/note="Site\_1: Salt; Site\_2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGACGCGCGCCGACATCGAGTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5'  
GAGAGAGATTCGATTCGATTAATTAATCCGCCCCCCC 3']. cDNA  
was cleaved with BamHI and XhoI. Vector: a modified  
pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 79 a 65 c 43 g 95 t

ORIGIN

Query Match 1.8%; Score 20; DB 9; Length 282;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
48 ctgcctcgtctcctactatg 67  
|||||  
24 ctgcctcgtctcctactatg 5

RESULT 6  
AA269349 339 bp mRNA linear EST 26-MAR-1997

LOCUS v090h06.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:748475

DEFINITION 5', mRNA sequence.

ACCESSION AA269349 GI:1907628

VERSION AA269349

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
1 (bases 1 to 339)

REFERENCE Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisels, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, R., Steptoe, M., Tan, F., Underwood, R., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.

TITLE The WashU-HMT Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maria M/Mouse EST Project  
WashU-HMT Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LMU; contact the  
IMAGE Consortium ([info@image.lml.gov](mailto:info@image.lml.gov)) for further information.  
MGI:457439

Putative full length read  
vector to vector length is 491  
Seq primer: -28ml3 rev2 RT from Amersham  
High quality sequence stop: 299.

FEATURES  
source Location/Qualifiers

1.339  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:748475"  
/clone\_1lb="Soares mouse NML"  
/tissue\_type="Liver"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCGAAGTGGAGCGCGCGGAGGATCTTTTCTTTTCTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 62 a 105 c 80 g 92 t

ORIGIN

Query Match 1.8%; Score 20; DB 9; Length 339;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
938 tggagctcaccctgactgc 957  
|||||  
153 tggagctcaccctgactgctc 172

RESULT 7  
AI834906/c 370 bp mRNA linear EST 14-JUL-1999

LOCUS AI834906

DEFINITION UI-M-AN1-afe-e-09-0-UI.s1 NIH\_BMAP\_MBG.N Mus musculus cDNA clone

ACCESSION AI834906

VERSION AI834906.1 GI:5469119

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
1 (bases 1 to 370)

REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477

TITLE Contact: Chin, H  
JOURNAL National Institute of Mental Health  
MEDLINE 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
COMMENT 20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mestr@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
normalized basal ganglia library cDNA library Preparation: M.B.  
Soares lab Clone distribution: NIH BMAP cDNA clones will be made  
available by the means that is soon to be determined. When NIH  
determines the means for distribution of the BMAP cDNA clones, this  
record will be updated accordingly when that means is determined.  
Seq primer: M13 Forward  
POLVA=yes.

FEATURES  
source Location/Qualifiers

1.370



```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-ANI-afe-e-09-0-UI"
/clone_lib="NIH-BMAP-MBG_N"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The
NIH-BMAP-MBG_N library is a normalized library constructed
from mouse basal ganglia. The tag is a string of 5
nucleotides present between the Not I site and the
oligo-dr track. The library was constructed as described
by Bonaldi, Lennon and Soares, Genome Research 6: 791-806
, 1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories.
TAG_LIB=NIH-BMAP-MBG_N
TAG_TISSUE=basal ganglia
TAG_SEQ=GTGAC"
BASE COUNT      89 a      84 c      99 g      98 t
ORIGIN

```

```

Query Match      1.8% Score 20; DB 9; Length 370;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 938 tgggctcacccctgattc 957
|||||
DB 366 TGGGCTTCACCTGATGTTTC 347

```

```

RESULT 8
AA590068      373 bp mRNA linear EST 16-SEP-1997
LOCUS      vnt7e06.r1 Barstead mouse proximal colon MRLRB6 Mus musculus cDNA
DEFINITION      clone IMAGE:1025314 5', mRNA sequence.
ACCESSION      AA590068
VERSION      AA590068.1 GI:2403448
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 373)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:576090
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 366.

```

## FEATURES

```

source
1. 373
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1025314"
/clone_lib="Barstead mouse proximal colon MRLRB6"
/dev_stage="7 day juvenile"
/lab_host="DH10B"

```

```

/notes="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCGATCTGAGTGGAGCGGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']. double-stranded cDNA was ligated to Eco RI adaptors
[ATTCGGAATTCCTTC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pRT3 vector.
Library constructed by Bob Barstead."
BASE COUNT      87 a      104 c      87 g      95 t
ORIGIN

```

```

Query Match      1.8% Score 20; DB 9; Length 373;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 938 tgggctcacccctgattc 957
|||||
DB 25 TGGGCTTCACCTGATGTTTC 44

```

```

RESULT 9
W30259      376 bp mRNA linear EST 11-SEP-1996
LOCUS      mc25a06.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
DEFINITION      IMAGE:349522 5', mRNA sequence.
ACCESSION      W30259
VERSION      W30259.1 GI:1310407
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 376)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:221322
Seq primer: EMP1mer
High quality sequence stop: 372.

```

## FEATURES

```

source
1. 376
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:349522"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Vector: pRT3D (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCGATCTGAGTGGAGCGGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']. double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pRT3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldi. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
BASE COUNT      60 a      116 c      103 g      97 t

```



Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:427434

Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 387.

## FEATURES

## SOURCE

1. .409

Location/Qualifiers

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:693874"

/clone\_lib="Soares mouse NML"

/tissue\_type="Liver"

/lab\_host="DHI0B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dt) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCGCAATCTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

## BASE COUNT

83 a 129 c 95 g 102 t

## ORIGIN

Query Match 1.8%; Score 20; DB 9; Length 409;  
Best Local Similarity 100.0%; Pred. No. 24;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 938 tgggctcaccctgattgc 957

DB 106 TGGCCTTCACCTGATGTTTC 125

## RESULT 13

## LOCUS

BF783504

DEFINITION 413 bp mRNA linear EST 12-JAN-2001  
602111279P1 NCL\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:4239571

5', mRNA sequence.

ACCESSION BF783504.1 GI:12088540

VERSION BF783504

KEYWORDS EST.

SOURCE house mouse.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM9852 row: n column: 20

High quality sequence stop: 413.

Location/Qualifiers

1. .413

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4239571"

/clone\_lib="NCI\_CGAP\_Kid14"

/lab\_host="DHI0B (T1 phage-resistant)"

/note="Organ: kidney. Vector: pCMV-SPOrt6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library. I"

## BASE COUNT

116 a 110 c 87 g 100 t

## ORIGIN

Query Match 1.8%; Score 20; DB 10; Length 413;  
Best Local Similarity 100.0%; Pred. No. 24;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 938 tgggctcaccctgattgc 957

DB 50 TGGCCTTCACCTGATGTTTC 69

## RESULT 14

## LOCUS

AA245797

DEFINITION 426 bp mRNA linear EST 10-MAR-1997  
mx06f04.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:679423

5', mRNA sequence.

ACCESSION AA245797.1 GI:1876618

VERSION AA245797

KEYWORDS EST.

SOURCE house mouse.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 426)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,D., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Thelings,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMNI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMNI Mouse EST Project

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:419127

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 425.

Location/Qualifiers

1. .426

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:679423"

/clone\_lib="Soares mouse NML"

/tissue\_type="Liver"

/lab\_host="DHI0B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dt) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCGCAATCTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 77 a 137 c 100 g 104 t 8 others

## ORIGIN

Query Match 1.8%; Score 20; DB 9; Length 426;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 938 tgggcttcacctgatgttc 957  
 |||  
 Db 202 TGGGCTTCACCCCTGATGTTC 221

Search completed: May 7, 2002, 17:25:00  
 Job time: 8130 sec

RESULT 15  
 AI042890 428 bp mRNA linear EST 01-JUL-1998  
 LOCUS uc77c05.x1 Sugano mouse liver mla Mus musculus cDNA clone  
 DEFINITION IMAGE:1431656.3', mRNA sequence.  
 ACCESSION AI042890  
 VERSION AI042890.1 GI:3286150  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 428)  
 Mairra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Maria M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:915724

Seq primer: custom primer used  
 High quality sequence stop: 351.

Location/Qualifiers  
 1. 428

#### FEATURES

##### source

/organism="Mus musculus"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1431656"  
 /clone\_lib="Sugano mouse liver mla"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: liver. Vector: pME18S-FL3. Site\_1: DraIII  
 (CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand CDNA  
 was primed with an oligo(dT) primer  
 [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded CDNA was  
 ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested  
 and cloned into distinct DraIII sites of the pME18S-FL3  
 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
 be used to isolate the cDNA insert. Size selection was  
 performed to exclude fragments <1.5kb. Library  
 constructed by Dr. Sumio Sugano (University of Tokyo  
 Institute of Medical Science). Custom primers for  
 sequencing: 5' end primer CTCTGCTCTAAAGCTGCG and 3' end  
 primer CGACTGCGACTGACACACA."

BASE COUNT 109 a 95 c 126 g 98 t  
 ORIGIN

Query Match 1.8%; Score 20; DB 9; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 938 tgggcttcacctgatgttc 957  
 |||  
 Db 363 TGGGCTTCACCCCTGATGTTC 344

---

